



105 18 33.3 492 1 TISD HUMAN  
 106 18 33.3 507 1 PGL2\_JUNAS  
 107 18 33.3 514 1 PGL2\_CHAOB  
 108 18 33.3 514 1 PGL2\_CRYJA  
 109 18 33.3 514 2 Q8H987  
 110 18 33.3 514 2 Q8H988  
 111 18 33.3 514 2 Q8H989  
 112 18 33.3 516 2 Q8AR84  
 113 18 33.3 522 2 Q8C8L5  
 114 18 33.3 523 2 Q8BQEL  
 115 18 33.3 525 1 MP11 RAT  
 116 18 33.3 555 2 Q8H139  
 117 18 33.3 584 1 CO8A HUMAN  
 118 18 33.3 632 2 Q7K6U7  
 119 18 33.3 634 2 Q91LA5  
 120 18 33.3 660 2 Q7QY47  
 121 18 33.3 667 2 Q9XTK3  
 122 18 33.3 671 2 Q9D678  
 123 18 33.3 675 1 KSC5 ECOLI  
 124 18 33.3 675 2 Q6KD00  
 125 18 33.3 675 2 Q8FDP9  
 126 18 33.3 690 2 Q7QE14  
 127 18 33.3 698 2 Q9D2L6  
 128 18 33.3 742 2 Q9N8X0  
 129 18 33.3 746 2 Q27526  
 130 18 33.3 750 2 P70921  
 131 18 33.3 762 2 Q7YVQ5  
 132 18 33.3 762 2 Q9E293  
 133 18 33.3 763 2 Q89K15  
 134 18 33.3 798 2 Q8C7A1  
 135 18 33.3 822 2 Q9V9C6  
 136 18 33.3 827 2 Q6TDN6  
 137 18 33.3 827 2 Q6TDP0  
 138 18 33.3 827 2 Q8CFM7  
 139 18 33.3 838 2 Q8R3S1  
 140 18 33.3 839 2 Q7R186  
 141 18 33.3 888 2 Q6KAT0  
 142 18 33.3 891 2 Q6TDN6  
 143 18 33.3 908 2 Q8OYR4  
 144 18 33.3 914 2 Q8CFM8  
 145 18 33.3 923 2 Q6TDN8  
 146 18 33.3 1015 2 Q6TDN7  
 147 18 33.3 1027 2 Q9FQS8  
 148 18 33.3 1046 2 Q612Q9  
 149 18 33.3 1048 2 Q05925  
 150 18 33.3 1144 2 Q6H8C5

## ALIGNMENTS

RESULT 1  
 YA40\_MYCPN STANDARD; PRT; 103 AA.  
 AC P75074;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DE 05-JUL-2004 (Rel. 44, Last annotation update)  
 DE Hypothetical protein MPN040 (B01\_orf103b).  
 GN OrderedLocusNames=MPN040; ORFNames=MP114;  
 OS Mycoplasma pneumoniae.  
 OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.  
 OX NCBI\_TaxID=2104;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 29342 / M129;  
 RX MEDLINE=97105885; PubMed=8948633; DOI=10.1093/nar/24.22.4420;  
 RA Himmelreich R., Hilbert H., Plegens H., Firkel E., Li B.-C.,  
 RA Herrmann R.;  
 RT "Complete sequence analysis of the genome of the bacterium Mycoplasma  
 RT pneumoniae";  
 RL Nucleic Acids Res. 24:4420-4449(1996).  
 CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL; AE000014; AB95762.1; -.  
 DR PIR; S73440; S73440.  
 KW Complete proteome; Hypothetical protein.  
 SQ SEQUENCE 103 AA; 11291 MW; 62E96884733993C6 CRC64;

Query Match 35.2%; Score 19; DB 1; Length 103;  
 Best Local Similarity 13.6%; Pred. No. 6.2e+02;  
 Matches 3; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

OY 3 CXXXXXXCXXXXXXCXXXXXXH 24  
 DB 53 CSSISFCSLASSAALRYSSSH 74

## RESULT 2

O7PER7 PRELIMINARY; PRT; 39 AA.  
 AC O7PER7;  
 DT 01-MAR-2004 (TRENBLrel. 26, Created)  
 DT 01-MAR-2004 (TRENBLrel. 26, Last sequence update)  
 DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)  
 DE ENSANGP0000023374.  
 GN Name=ENSANGG0000020343;  
 OS Anopheles gambiae str. PEST.  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Anopheles.  
 OX NCBI\_TaxID=180454;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=PEST;  
 RA Anopheles Genome Sequencing Consortium;  
 RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.  
 CC -!- CAUTION: The sequence shown here is derived from an  
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is  
 CC preliminary data.  
 DR EMBL; AA801008326; EAA45619.1; -.  
 SQ SEQUENCE 39 AA; 4317 MW; 6963C9D228D7C8C7 CRC64;

Query Match 33.3%; Score 18; DB 2; Length 39;  
 Best Local Similarity 28.6%; Pred. No. 6.5e+02;  
 Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 3 CXXXXXC 9  
 DB 2 CAATASC 8

## RESULT 3

O7UWS9 PRELIMINARY; PRT; 49 AA.  
 AC O7UWS9;  
 DT 01-OCT-2003 (TRENBLrel. 25, Created)  
 DT 01-OCT-2003 (TRENBLrel. 25, Last sequence update)  
 DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)  
 DE Hypothetical protein.  
 GN OrderedLocusNames=RB1825;  
 OS Rhodospirillum rubrum.  
 OC Bacteria; Planctomycetes; Planctomycetacia; Planctomycetales;  
 OC Planctomycetaceae; Pirellula.  
 OX NCBI\_TaxID=117;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=1;  
 RX MEDLINE=22735913; PubMed=12835416; DOI=10.1073/pnas.1431443100;  
 RA Gloeckner F.O., Kube M., Bauer M., Teeling H., Lombardot T.,

RA Ludwig W., Gade D., Beck A., Borzym K., Heitmann K., Rabus R.,  
 RA Schlesner H., Amann R., Reinhardt R.;  
 RT "Complete genome sequence of the marine planctomycete *Pirellula* sp.  
 RT strain 1.";   
 RL Proc. Natl. Acad. Sci. U.S.A. 100:8298-8303 (2003).  
 DR EMBL; BX294133; CAD72283.1; --  
 KW Complete proteome; Hypothetical protein.  
 SQ SEQUENCE 49 AA; 5561 MW; D47FDS9BFF7BA40 CRC64;

Query Match 33.3%; Score 18; DB 2; Length 49;  
 Best Local Similarity 28.6%; Pred. No. 7.5e+02; Indels 5; Gaps 0;  
 Matches 2; Conservative 0; Mismatches 5;

Qy 3 CXXXXXC 9  
 |  
 Db 3 CXXXXXC 9

## RESULT 4

Q7PE97 PRELIMINARY; PRT; 58 AA.  
 AC Q7PE97;  
 DT 01-MAR-2004 (TrEMBLrel. 26, Created)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
 DE ENSANGP0000022719 (Fragment).  
 GN Name=ENSANG00000221658;  
 OS Anopheles gambiae str. PEST.  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Anopheles.  
 OX NCBI\_TaxID=180454;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=PEST;  
 RA Anopheles Genome Sequencing Consortium;  
 RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.  
 CC -! CAUTION: The sequence shown here is derived from an  
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is  
 CC preliminary data.  
 DR EMBL; AAB01004888; EAA45826.1; --  
 FT NON\_TER 1 58  
 FT NON\_TER 58  
 SQ SEQUENCE 58 AA; 6678 MW; BDD2ESD73967DAB5 CRC64;

Query Match 33.3%; Score 18; DB 2; Length 58;  
 Best Local Similarity 28.6%; Pred. No. 8.4e+02; Indels 5; Gaps 0;  
 Matches 2; Conservative 0; Mismatches 5;

Qy 3 CXXXXXC 9  
 |  
 Db 28 CAATASC 34

## RESULT 5

Q86YX3 PRELIMINARY; PRT; 61 AA.  
 AC Q86YX3;  
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
 DE MTE.  
 GN Name=MTE;  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Yu L.;  
 RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.

CC -! FUNCTION: Metallothioneins have a high content of cysteine  
 CC residues that bind various heavy metals (By similarity).  
 CC -! SIMILARITY: Belongs to the metallothionein superfamily. Family 1.

DR EMBL; AF348997; AAO32957.1; --  
 DR HSP; P02795; 1MHU.  
 DR GO; GO:0046872; F:metal ion binding; IEA.  
 DR InterPro; IPR003019; Metallothion\_1.  
 DR InterPro; IPR000006; Metallothion\_1.  
 DR Pfam; PF00131; Metallothio; 1.  
 DR PRINTS; PR00860; MTVERTEBRATE.  
 KW Metal-binding; Metal-thiolate cluster.  
 SQ SEQUENCE 61 AA; 6165 MW; 7F71CE7D1C37762A CRC64;

Query Match 33.3%; Score 18; DB 2; Length 61;  
 Best Local Similarity 28.6%; Pred. No. 8.7e+02; Indels 5; Gaps 0;  
 Matches 2; Conservative 0; Mismatches 5;

Qy 3 CXXXXXC 9  
 |  
 Db 7 CSTSSSC 13

## RESULT 6

CMC1\_EIMAC STANDARD; PRT; 79 AA.  
 AC P21959;  
 DT 01-AUG-1991 (Rel. 19, Created)  
 DT 01-AUG-1991 (Rel. 19, Last sequence update)  
 DT 05-JUL-2004 (Rel. 44, Last annotation update)  
 DE EAMZP30-47 protein (Fragment).  
 GN Name=CMC17;  
 OS Eimeria acervulina.  
 OC Eukaryota; Alveolata; Apicomplexa; Coccidia; Eimeriida; Eimeriidae;  
 OC Eimeria.  
 OX NCBI\_TaxID=5801;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=9018433; PubMed=1690144; DOI=10.1016/0014-4894(90)90117-U;  
 RA Jenkins M.C., Lillehoj H.S., Barta J.R., Danforth H.D.,  
 RA Strohllein D.A.;  
 RT "Eimeria acervulina: cloning of a cDNA encoding an immunogenic region  
 RT of several related merozoite surface and rhoptry proteins.";  
 RL Exp. Parasitol. 70:353-362(1990).  
 CC -! SUBCELLULAR LOCATION: Surface membrane and internal rhoptries.  
 CC -! DEVELOPMENTAL STAGE: Merozoite.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; M37843; AAA62796.1; --  
 KW Antigen; Membrane; Merozoite.  
 FT NON\_TER 1 79  
 FT NON\_TER 79  
 SQ SEQUENCE 79 AA; 8703 MW; 0B2B6CDB65FB4330 CRC64;

Query Match 33.3%; Score 18; DB 1; Length 79;  
 Best Local Similarity 28.6%; Pred. No. 1e+03; Indels 5; Gaps 0;  
 Matches 2; Conservative 0; Mismatches 5;

Qy 3 CXXXXXC 9  
 |  
 Db 30 CXXXXXC 36

## RESULT 7

Q7PE39 PRELIMINARY; PRT; 87 AA.  
 AC Q7PE39;  
 DT 01-MAR-2004 (TrEMBLrel. 26, Created)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)

DE ENSANGP0000023751 (Fragment).  
 GN Name=ENSANGG0000020038  
 OS Anopheles gambiae str. FEST.  
 OC Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Anopheles.  
 OX NCBI\_TaxID=180454;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=PEST;  
 RA Anopheles Genome Sequencing Consortium;  
 RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.  
 CC -!- CAUTION: The sequence shown here is derived from an  
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is  
 CC preliminary data.  
 DR EMBL; AAAB01002921; EAA45886.1; --  
 FT NON TER 1  
 FT NON TER 87  
 SQ SEQUENCE 87 AA; 10053 MW; 45F57BD9A1AE0FC3 CRC64;  
 Query Match 33.3%; Score 18; DB 2; Length 87;  
 Best Local Similarity 28.6%; Pred. No. 1.1e+03;  
 Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
 QY 3 CXXXXXC 9  
 Db 28 CAATASC 34  
 RESULT 8  
 Q94AZ8 PRELIMINARY; PRT; 88 AA.  
 ID Q94AZ8  
 AC Q94AZ8  
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)  
 DE AT422230/T10114 60 (Hypothetical protein).  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.  
 OX NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Koesema E., Chen H., Cheuk R., Kim C.J., Meyers M.C., Shinn P.,  
 RA Banh J., Bowser L., Carninci P., Dale J.M., Gibson H.A.,  
 RA Goldsmith A.D., Hayashizaki Y., Ishida J., Jiang P.X., Jones T.,  
 RA Kamiya A., Karlin-Neumann G., Kawai J., Lam B., Lee J.M., Lin J.,  
 RA Liu S.X., Miranda M., Narusaka M., Nguyen M., Onodera C.S., Palm C.J.,  
 RA Pham P.K., Quach H.L., Sakurai T., Satou M., Seki M., Southwick A.,  
 RA Tang C.C., Toriumi M., Yamada K., Yu G., Yu S., Shinozaki K.,  
 RA Davis R.W., Theologis A., Ecker J.R.;  
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22088475; PubMed=12093376;  
 RA Haas B.J., Volforsky N., Town C.D., Troukhan M., Alexandrov N.,  
 RA Feldmann K.A., Flavell R.B., White O., Salzberg S.L.;  
 RT "Full-length messenger RNA sequences greatly improve genome  
 RT annotation."  
 RL Genome Biol. 3:RESEARCH0029-RESEARCH0029(2002).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RA Brover V., Troukhan M., Alexandrov N., Lu Y.-P., Flavell R.,  
 RA Feldmann K.;  
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RA Cheuk R., Chen H., Kim C.J., Meyers M.C., Shinn P., Banh J.,  
 RA Bowser L., Carninci P., Chang E., Dale J.M., Goldsmith A.D.,  
 RA Hayashizaki Y., Ishida J., Jones T., Kamiya A., Karlin-Neumann G.,  
 RA Kawai J., Lam B., Lee J.M., Lin J., Miranda M., Narusaka M.,  
 RA Nguyen M., Onodera C.S., Palm C.J., Quach H.L., Sakurai T., Satou M.,  
 RA Seki M., Southwick A., Tang C.C., Toriumi M., Wu H.C., Yamada K.,

RA Yamamura Y., Yu G., Yu S., Shinozaki K., Davis R.W., Theologis A.,  
 RA Ecker J.R.;  
 RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AY045581; AAK73939.1; --  
 DR EMBL; AY084624; AAM61187.1; --  
 DR EMBL; AY094030; AAM16186.1; --  
 DR GO; GO:0006952; P:defense response; IEA.  
 DR InterPro; IPR003614; Knott1.  
 DR SMART; SM00505; Knott1; 1.  
 KW Hypothetical protein.  
 SQ SEQUENCE 88 AA; 9071 MW; 2345DAB2B66E96A8 CRC64;  
 Query Match 33.3%; Score 18; DB 2; Length 88;  
 Best Local Similarity 28.6%; Pred. No. 1.1e+03;  
 Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
 QY 3 CXXXXXC 9  
 Db 40 CTSSSTC 46  
 RESULT 9  
 Q63316 PRELIMINARY; PRT; 111 AA.  
 ID Q63316  
 AC Q63316  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
 DE Ly6-A antigen (Fragment).  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX STRAIN=Sprague-Dawley; TISSUE=Kidney;  
 RX MEDLINE=90152758; PubMed=2154400;  
 RA Friedman S., Palfree R.G.E., Sirlin S., Haemmerling U.;  
 RT "Analysis of three distinct Ly6-A-related cDNA sequences isolated from  
 RT rat kidney."  
 RL Immunogenetics 31:104-111(1990).  
 DR EMBL; M30692; AAA41545.1; --  
 DR HSSP; Q9Y111; 1HC9.  
 DR InterPro; IPR001526; LY6 UPAR.  
 DR Pfam; PF00021; UPAR\_LY6\_1.  
 DR SMART; SM00134; LU; 1.  
 FT NON TER 111  
 SQ SEQUENCE 111 AA; 11755 MW; B0A42D1B3ED148C4 CRC64;  
 Query Match 33.3%; Score 18; DB 2; Length 111;  
 Best Local Similarity 28.6%; Pred. No. 1.3e+03;  
 Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
 QY 3 CXXXXXC 9  
 Db 17 CSSTATC 23  
 RESULT 10  
 Q6K499 PRELIMINARY; PRT; 113 AA.  
 ID Q6K499  
 AC Q6K499;  
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)  
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)  
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)  
 DE Hypothetical protein QJ1595\_D08.20.  
 GN Name=QJ1595\_D08.20;  
 OS Oryza sativa (japonica cultivar-group).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC Ehrhartoideae; Magnoliaceae; Oryza.  
 OX NCBI\_TaxID=39947;  
 RN [1]

RP SEQUENCE FROM N.A.  
 RA Sasaki T., Matsumoto T., Hattori M., Sakaki Y., Katayose Y.;  
 RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AP005574; BAD22270.1; --  
 KW Hypothetical protein.  
 SQ SEQUENCE 113 AA; 12326 MW; C2A6C98129828244 CRC64;

Query Match 33.3%; Score 18; DB 2; Length 113;  
 Best Local Similarity 28.6%; Pred. No. 1.3e+03;  
 Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 CXXXXXC 9  
 DB 35 CSASACA 41

RESULT 11  
 ID O40635 PRELIMINARY; PRT; 116 AA.  
 AC O40635;  
 DT 01-JAN-1998 (TREMBLrel. 05, Created)  
 DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)  
 DT 05-JUL-2004 (TREMBLrel. 27, Last annotation update)  
 DE CD59 protein (Viral CD59 antigen).  
 GN Name=orf15; Synonyms=VCD59;  
 OS Saimiriine herpesvirus 2.  
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;  
 OC Gammaherpesvirinae; Rhadinovirus.  
 OX NCBI\_TaxID=10381;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C-488;  
 RX MEDLINE=98037620; PubMed=9371569;  
 RA Knappe A., Hiller C., Thurauf M., Wittmann S., Hofmann H.,  
 RA Fleckenstein B., Fickenscher H.;  
 RT "The superantigen-homologous viral immediate-early gene iel4/vseg in  
 RT herpesvirus saimiri-transformed human T cells.";  
 RL J. Virol. 71:9124-9133(1997).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C488;  
 RX MEDLINE=22918177; PubMed=14554077; DOI=10.1016/S0042-6822(03)00449-5;  
 RA Ensser A., Thurauf M., Wittmann S., Fickenscher H.;  
 RT "The genome of herpesvirus saimiri C488 which is capable of  
 RT transforming human T cells.";  
 RL Virology 314:471-487(2003).  
 DR EMBL; Y13183; CAA73629.1; --  
 DR EMBL; AJ410493; CAC94310.1; --  
 DR HSSP; P13987; IERG.  
 DR InterPro; IPR003632; LY-6\_CD59.  
 DR InterPro; IPR001526; LY6\_UPAR.  
 DR Pfam; PF00021; UPAR\_LY6\_1.  
 DR ProDom; PD003128; LY-6\_CD59; 1.  
 DR SMART; SM00134; LU; 1.  
 DR PROSITE; PS00983; LY6\_UPAR; 1.  
 KW Glycoprotein; GPI-anchor; Lipoprotein; Signal.  
 FT SIGNAL 1 19 Potential.  
 FT CHAIN 20 96 Surface glycoprotein CD59 homolog.  
 FT PROPEP 97 121 Removed in mature form (Potential).  
 FT DOMAIN 20 104 UPAR/LY6.  
 FT DISULFID 22 45 By similarity.  
 FT DISULFID 25 32 By similarity.  
 FT DISULFID 38 58 By similarity.  
 FT DISULFID 64 82 By similarity.  
 FT DISULFID 83 88 By similarity.  
 FT LIPID 96 96 GPI-anchor amidated asparagine (by host)  
 FT CARBOHYD 24 24 N-linked (GlcNAc...) (Potential).  
 FT SEQUENCE 116 AA; 13172 MW; 68E35CE923FF189 CRC64;

Query Match 33.3%; Score 18; DB 2; Length 116;  
 Best Local Similarity 28.6%; Pred. No. 1.3e+03;  
 Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 CXXXXXC 9  
 DB 32 CTTSTSC 38

RESULT 12  
 ID CD59\_SHV21 STANDARD; PRT; 121 AA.  
 AC Q00986;  
 DT 01-APR-1993 (Rel. 25, Created)  
 DT 01-APR-1993 (Rel. 25, Last sequence update)  
 DT 05-JUL-2004 (Rel. 44, Last annotation update)

DE Surface glycoprotein CD59 homolog precursor.  
 GN Name=15;  
 OS Saimiriine herpesvirus 2 (strain 11) (SAHV-2) (Herpesvirus saimiri).  
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;  
 OC Gammaherpesvirinae; Rhadinovirus.  
 OX NCBI\_TaxID=10383;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=92333688; PubMed=1321287;  
 RA Albrecht J.-C., Nicholas J., Biller D., Cameron K.R., Biesinger B.,  
 RA Newman C., Wittmann S., Craxton M.A., Coleman H., Fleckenstein B.,  
 RA Honess R.W.;  
 RT "Primary structure of the herpesvirus saimiri genome.";  
 RL J. Virol. 66:5047-5058(1992).  
 RN [2]  
 RP SIMILARITY TO CD59.  
 RX MEDLINE=92410640; PubMed=1382344;  
 RA Albrecht J.-C., Nicolas J., Cameron K.R., Newman C., Fleckenstein B.,  
 RA Honess R.W.;  
 RT "Herpesvirus saimiri has a gene specifying a homologue of the cellular  
 RT membrane glycoprotein CD59.";  
 RL Virology 190:527-530(1992).  
 CC -|- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor (By  
 CC similarity).  
 CC -|- SIMILARITY: Contains 1 UPAR/LY6 domain.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; X64346; CAA45638.1; --  
 DR EMBL; X64273; CAA45565.1; --  
 DR HSSP; P13987; IERG.  
 DR InterPro; IPR003632; LY-6\_CD59.  
 DR InterPro; IPR001526; LY6\_UPAR.  
 DR Pfam; PF00021; UPAR\_LY6\_1.  
 DR ProDom; PD003128; LY-6\_CD59; 1.  
 DR SMART; SM00134; LU; 1.  
 DR PROSITE; PS00983; LY6\_UPAR; 1.  
 KW Glycoprotein; GPI-anchor; Lipoprotein; Signal.  
 FT SIGNAL 1 19 Potential.  
 FT CHAIN 20 96 Surface glycoprotein CD59 homolog.  
 FT PROPEP 97 121 Removed in mature form (Potential).  
 FT DOMAIN 20 104 UPAR/LY6.  
 FT DISULFID 22 45 By similarity.  
 FT DISULFID 25 32 By similarity.  
 FT DISULFID 38 58 By similarity.  
 FT DISULFID 64 82 By similarity.  
 FT DISULFID 83 88 By similarity.  
 FT LIPID 96 96 GPI-anchor amidated asparagine (by host)  
 FT CARBOHYD 24 24 N-linked (GlcNAc...) (Potential).  
 FT SEQUENCE 121 AA; 13814 MW; 56A8BDC07B953AA3 CRC64;

Query Match 33.3%; Score 18; DB 1; Length 121;  
 Best Local Similarity 28.6%; Pred. No. 1.3e+03;  
 Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 CXXXXXC 9  
 DB 32 CTTSTSC 38

RESULT 13  
 ID Q9GM28 PRELIMINARY; PRT; 124 AA.  
 AC Q9GM28;  
 DT 01-MAR-2001 (TREMBLrel. 16, Created)  
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)

```

DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecinae; Macaca.
OX NCBI_TaxID=9541;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain parietal lobe;
RX MEDLINE=2145851; PubMed=11574149; DOI=10.1016/S0378-1119(01)00665-5;
RA Oseada N., Hida M., Kusuda J., Tanuma R., Iseki K., Hirata M., Suto Y.,
RA Hirai M., Terao K., Suzuki Y., Sugano S., Hashimoto K.;
RT "Assignment of 118 novel cDNAs of cynomolgus monkey brain to human
chromosomes.";
RL Gene 275:31-37(2001).
DR EMBL; AB049858; BAB16744.1; -.
KW Hypothetical protein.
SQ SEQUENCE 124 AA; 13904 MW; B87262632AF6D149 CRC64;
Query Match 33.3%; Score 18; DB 2; Length 124;
Best Local Similarity 28.6%; Pred. No. 1.4e+03;
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 3 CXXXXXC 9
DB 16 CAASSTC 22

RESULT 14
QID2 PRELIMINARY; PRT; 129 AA.
AC Q6ILG2;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE HDC09485.
GN ORFNames=HDC09485;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC PubMed=14709175; DOI=10.1186/gb-2003-5-1-r3;
RA Hild M., Beckmann B., Haas S., Koch B., Solovyev V., Busold C.,
RA Fellenberg K., Boutros M., Vingron M., Sauer F., Hohlseil J., Paro R.;
RT "An integrated gene annotation and transcriptional profiling approach
towards the full gene content of the Drosophila genome.";
RL Genome Biol. 5:R3-R3(2003).
CC -!- MISCELLANEOUS; The sequence shown here is derived from an
CC EMBL/GenBank/DBJ third party annotation (TPA) entry.
DR EMBL; BK002054; DRA02899.1; -.
SQ SEQUENCE 129 AA; 13635 MW; 1E5850AE37A3F647 CRC64;
Query Match 33.3%; Score 18; DB 2; Length 129;
Best Local Similarity 28.6%; Pred. No. 1.4e+03;
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 3 CXXXXXC 9
DB 91 CSSTSC 97

RESULT 15
Q8N7Z3 PRELIMINARY; PRT; 131 AA.
AC Q8N7Z3;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Hypothetical protein FLJ40194.

```

```

OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RX PubMed=14702039; DOI=10.1038/ng1285;
RA Ota T., Suzuki Y., Nishikawa T., Ohtsuki T., Sugiyama T., Irie R.,
RA Wakamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H.,
RA Sekine M., Ohtsuka M., Nishi T., Shibahara T., Tanaka T., Ishii S.,
RA Yamamoto K., Saito K., Kawai Y., Isono Y., Nakamura Y., Nagahara K.,
RA Murakami K., Yasuda T., Iwayanagi T., Wagatsuma M., Shiratori A.,
RA Sudo H., Hosoiri T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
RA Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa B., Omura Y.,
RA Abe K., Kamihara K., Katsuta N., Sato K., Tanikawa M., Yamazaki M.,
RA Ninomiya K., Ishibashi T., Yamashita H., Murakawa K., Fujimori K.,
RA Tanai H., Kimata M., Watanabe M., Hirakawa S., Chiba Y., Ishida S.,
RA Ono Y., Takiguchi S., Watanabe S., Yosida M., Hotuta T., Kusano J.,
RA Kanehori K., Takahashi-Fujii A., Hara H., Tanase T., Nomura Y.,
RA Togiya S., Komai F., Hara R., Takeuchi K., Arita M., Ohtsuka S.,
RA Yashikawa Y., Matsunawa H., Ichihara T., Shionhara N., Sano S.,
RA Moriya S., Momiyama H., Satoh N., Takami S., Terashima Y., Suzuki O.,
RA Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H.,
RA Hishigaki H., Watanabe T., Sugiyama A., Takemoto M., Kawakami B.,
RA Yamazaki M., Watanabe K., Kumagai A., Itakura S., Fukuzumi Y.,
RA Fujimori Y., Komiyama M., Tashiro H., Tanigami A., Fujiwara T.,
RA Ono T., Yamada K., Fujii Y., Ozaki K., Hiroo M., Ohmori Y.,
RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
RA Okitani R., Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T.,
RA Matsumura K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M.,
RA Togashi T., Oyama M., Hata H., Watanabe M., Komatsu T.,
RA Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi Y., Nakagawa K.,
RA Okumura K., Nagase T., Nomura N., Kikuchi H., Masuho Y., Yamashita R.,
RA Nakai K., Yada T., Nakamura Y., Ohara O., Isogai T., Sugano S.;
RT "Complete sequencing and characterization of 21,243 full-length human
cDNAs.";
RL Nat. Genet. 36:40-45(2004).
DR EMBL; AK037513; BAC05083.1; -.
SQ SEQUENCE 131 AA; 14426 MW; 39A5B4F5292B2B2D CRC64;
Query Match 33.3%; Score 18; DB 2; Length 131;
Best Local Similarity 28.6%; Pred. No. 1.4e+03;
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 3 CXXXXXC 9
DB 113 CTATTAC 119

RESULT 16
Q63317 PRELIMINARY; PRT; 135 AA.
ID Q63317;
AC Q63317;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Rat Ly6-B antigen.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90152758; PubMed=2154400;
RA Friedman S., Palfrey R.G.E., Sirin S., Haemmerling U.;
RT "Analysis of three distinct Ly6-A-related cDNA sequences isolated from
rat kidney.";
RL Immunogenetics 31:104-111(1990).
DR EMBL; M30689; AAA1546.1; -.
DR PIR; A45835; A45835.
DR HSSP; Q9Y111; 2ABX.

```

DR InterPro; IPR001526; LY6 UPAR.  
 DR Pfam; PF00021; UPAR\_LY6; 1.  
 DR SMART; SM00134; LU; 1.  
 SQ SEQUENCE 135 AA; 14130 MW; E4D6ADC56B61A053 CRC64;

Query Match 33.3%; Score 18; DB 2; Length 135;  
 Best Local Similarity 28.6%; Pred. No. 1.4e+03;  
 Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 3 CXXXXXC 9  
 Db 41 CSSTATC 47

RESULT 17  
 Q24987 PRELIMINARY; PRT; 137 AA.  
 AC Q24987; 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Surface protein (Fragment).  
 GN Name=tp; Giardia lamblia (Giardia intestinalis).  
 OS Eukaryota; Diplomonadida; Hexamitidae; Giardia.  
 OX NCBI\_TaxID=5741;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA EY P.L., Darby J.M., Mayrhofer G.;  
 RL Submitted (JUN-1993) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; L16973; AAB59198.1; -.  
 DR InterPro; IPR005127; Giardia\_VSP.  
 DR InterPro; IPR009030; Grow\_fac\_recept.  
 DR Pfam; PF03302; VSP; 1.  
 FT NON\_TER 1  
 FT NON\_TER 137  
 SQ SEQUENCE 137 AA; 14148 MW; A340072153005D58 CRC64;

Query Match 33.3%; Score 18; DB 2; Length 137;  
 Best Local Similarity 28.6%; Pred. No. 1.4e+03;  
 Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 3 CXXXXXC 9  
 Db 103 CTTSSAC 109

RESULT 18  
 Q8C2X1 PRELIMINARY; PRT; 138 AA.  
 AC Q8C2X1; 01-MAR-2003 (TrEMBLrel. 23, Created)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
 DE Mus musculus 2 days pregnant adult female ovary cDNA, RIKEN full-length enriched library, clone:E330019L22 product:hypothetical protein, full insert sequence.  
 DE Mus musculus (Mouse).  
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Ovary;  
 RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6875(99)03004-9;  
 RA Carninci P., Hayashizaki Y.;  
 RT "High-efficiency full-length cDNA cloning."; Meth. Enzymol. 303:19-44(1999).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Ovary;  
 RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;  
 RA RIKEN FANTOM Consortium;

RT "Functional annotation of a full-length mouse cDNA collection."; Nature 409:685-690(2001).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Ovary;  
 RA The FANTOM Consortium;  
 RT "Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs."; Nature 420:563-573(2002).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Ovary;  
 RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;  
 RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M., Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;  
 RT "Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes."; Genome Res. 10:1617-1630(2000).  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Ovary;  
 RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;  
 RA Shibata K., Itoh M., Aizawa K., Nagao S., Sasaki N., Carninci P., Konno H., Akiyama J., Nishi K., Kitsunai T., Tashiro H., Itoh M., Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A., Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K., Fujiwaki S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M., Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J., Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;  
 RT "RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer."; Genome Res. 10:1757-1771(2000).  
 RN [6]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Ovary;  
 RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P., Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W., Hayashida K., Hayatsu N., Hiramoto K., Hirooka T., Hirozane T., Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T., Katoh K., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S., Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M., Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y., Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H., Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M., Tagawa A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;  
 RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AK087776; BAC39998.1; -.  
 SQ SEQUENCE 138 AA; 14564 MW; 42565D90452C4749 CRC64;

Query Match 33.3%; Score 18; DB 2; Length 138;  
 Best Local Similarity 28.6%; Pred. No. 1.5e+03;  
 Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 3 CXXXXXC 9  
 Db 90 CSSSSAC 96

RESULT 19  
 Q6LBB8 PRELIMINARY; PRT; 139 AA.  
 AC Q6LBB8; 05-JUL-2004 (TrEMBLrel. 27, Created)  
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)  
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)  
 DE Hypothetical protein.  
 OS Oligotropha carboxidovorans (Pseudomonas carboxydovorans).  
 OG Plasmid pHC3.  
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales; Bradyrhizobiales; Oligotropha.

OX NCBI\_TaxID=40137;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=OMS;  
 RX MEDLINE=95238294; PubMed=7721710;  
 RA Schuebel U., Kraut M., Moersdorf G., Meyer O.;  
 RT "Molecular characterization of the gene cluster coxMSL encoding the  
 RT mycobacterium-containing carbon monoxide dehydrogenase of *Oligotropha*  
 RT carboxidovorans".  
 RL J. Bacteriol. 177:2197-2197(1995).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=OMS;  
 RX MEDLINE=97464431; PubMed=9324252;  
 RA Santiago B., Meyer O.;  
 RT "Purification and molecular characterization of the H2 uptake  
 RT membrane-bound NiFe-hydrogenase from the carboxidotrophic bacterium  
 RT *Oligotropha carboxidovorans*".  
 RL J. Bacteriol. 179:6053-6060(1997).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=OMS;  
 RA Santiago B., Schuebel U., Egelseer C., Meyer O.;  
 RT "Sequence analysis, characterization and CO-specific transcription of  
 RT the cox gene cluster on the megaplasmid pHC3 of *Oligotropha*  
 RT carboxidovorans".  
 RL Gene 236:1157-1247(1999).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=OMS;  
 RX PubMed=14644498; DOI=10.1016/j.gene.2003.08.027;  
 RA Fuhrmann S., Ferner M., Jeffke T., Henne A., Gottschalk G., Meyer O.;  
 RT "Complete nucleotide sequence of the self-transmissible circular  
 RT megaplasmid pHC3 of *Oligotropha carboxidovorans*: Function in the  
 RT chemolithoautotrophic utilization of CO, H2 and CO2".  
 RL Gene 322:67-75(2003).  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=OMS;  
 RA Schuebel U.;  
 RT Submitted (JUL-1993) to the EMBL/GenBank/DBJ databases.  
 RN [6]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=OMS;  
 RA Fuhrmann S.;  
 RT Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.  
 RL EMBL; X82447; CAG28441.1; -;  
 SQ SEQUENCE 139 AA; 15501 MW; F67DC09839ABFA6A CRC64;  
 Query Match 33.3%; Score 18; DB 2; Length 139;  
 Best Local Similarity 28.6%; Pred. No. 1.5e+03;  
 Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
 QY 3 CXXXXX 9  
 DB 103 CASSTSC 109  
 RESULT 20  
 Q96AC2  
 ID Q96AC2 PRELIMINARY; PRT; 141 AA.  
 AC Q96AC2;  
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
 DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)  
 DE Hypothetical protein FLJ37182 (PSCA Hlog).  
 GN ORFNames=UNQ3079;  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]

RP SEQUENCE FROM N.A.  
 RC TISSUE=Pancreas;  
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hong F.,  
 RA Datschenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
 RA Krzywinski M.I., Skalska U., Smallos D.E., Schnerch A., Schein J.E.,  
 RA Jones S.J., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences".  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Pancreas;  
 RA Strausberg R.;  
 RT Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Alzheimer cortex;  
 RX PubMed=14702039; DOI=10.1038/ng1285;  
 RA Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,  
 RA Wakamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H.,  
 RA Sekine M., Obayashi M., Nishi T., Shibahara T., Tanaka T., Ishii S.,  
 RA Yamamoto J., Saito K., Kawai Y., Isono Y., Nakamura Y., Nagahari K.,  
 RA Murakami K., Yasuda T., Iwayanagi T., Wagatsuma M., Shiratori K.,  
 RA Sudo H., Hosoiri T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,  
 RA Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa E., Omura Y.,  
 RA Abe K., Kamihara K., Katsuta N., Sato K., Tanikawa M., Yamazaki M.,  
 RA Niimi H., Kimata M., Watanabe M., Hiraoka S., Chiba Y., Ishida S.,  
 RA Tanai H., Kimata M., Watanabe M., Hiraoka S., Chiba Y., Ishida S.,  
 RA Ono Y., Takiguchi S., Watanabe S., Yosida M., Hotuta T., Kusano J.,  
 RA Kanehori K., Takahashi-Fujii A., Hara H., Tanase T., Nomura Y.,  
 RA Togiya S., Komai F., Hara R., Takeuchi K., Arita M., Imose N.,  
 RA Musashino K., Yuuki H., Ooshima A., Sasaki N., Aotsuka S.,  
 RA Yoshikawa Y., Matsunawa H., Ichihara T., Shiohata N., Sano S.,  
 RA Moriya S., Momiyama H., Satoh N., Takami S., Terashima Y., Suzuki O.,  
 RA Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H.,  
 RA Hishigaki H., Watanabe T., Sugiyama A., Takemoto M., Kawakami B.,  
 RA Yamazaki M., Watanabe K., Kumagai A., Itakura S., Fukuzumi Y.,  
 RA Fujimori Y., Komiyama M., Tashiro H., Tanigami A., Fujiwara T.,  
 RA Ono T., Yamada K., Fujii Y., Ozaki K., Hirao M., Ohmori Y., Okamoto S.,  
 RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Senba T.,  
 RA Okitani R., Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T.,  
 RA Matsumura K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M.,  
 RA Togashi T., Oyama M., Hata H., Watanabe M., Komatsu T.,  
 RA Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi Y., Nakagawa K.,  
 RA Okumura K., Nagase T., Nomura N., Kikuchi H., Masuho Y., Yanashita R.,  
 RA Nakai K., Yada T., Nakamura Y., Ohara O., Isogai T., Sugano S.;  
 RT "Complete sequencing and characterization of 21,243 full-length human  
 RT cDNAs".  
 RL Nat. Genet. 36:40-45(2004).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22887296; PubMed=12975309; DOI=10.1101/gr.1293003;  
 RA Clark H.F., Gurney A.L., Abaya E., Baker K., Baldwin D., Brush J.,  
 RA Chen J., Chow B., Chui C., Crowley C., Currell B., Deuel B., Dowd P.,  
 RA Eaton D., Foster J., Grimaldi C., Gu Q., Hass P.E., Heldens S.,  
 RA Huang A., Kim H.S., Klimowski L., Jin Y., Johnson S., Lee J.,  
 RA Lewis L., Liao D., Mark M., Robbie E., Sanchez C., Schoenfeld J.,  
 RA Seshagiri S., Simmons L., Singh J., Smith V., Stinson J., Vasta A.,  
 RA Vandlen R., Watanabe C., Wieand D., Woodde K., Xie M.H., Yaneura D.,



RA Yi S., Yu G., Yuan J., Zhang M., Zhang Z., Goddard A., Wood W.I.,  
RA Godowski P.;  
RT "The secreted protein discovery initiative (SPDI), a large-scale  
RT effort to identify novel human secreted and transmembrane proteins: a  
RT bioinformatics assessment.";  
RL Genome Res. 13:2265-2270(2003).  
DR EMBL; BC017318; RAH17318.1; -;  
DR EMBL; AK094501; BAC04368.1; -;  
DR EMBL; AY358628; AAQ89991.1; -;  
KW Hypothetical protein.  
SQ SEQUENCE 141 AA; 15240 MW; CFA6D98BA90792E3 CRC64;  
Query Match 33.3%; Score 18; DB 2; Length 141;  
Best Local Similarity 28.6%; Pred. No. 1.5e+03;  
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
OY 3 CXXXXXX 9  
DB 71 CASSAAC 77  
RESULT 21  
ID Q66H42 PRELIMINARY; PRT; 141 AA.  
AC Q66H42;  
DT 25-OCT-2004 (TREMBLrel. 28, Created)  
DT 25-OCT-2004 (TREMBLrel. 28, Last sequence update)  
DT 25-OCT-2004 (TREMBLrel. 28, Last annotation update)  
DE Hypothetical protein.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Kidney;  
RX PubMed=12477932; DOI=10.1073/pnas.242603899;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P.,  
RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
RA Krywinski M.I., Skalska U., Smalusz D.E., Schnerch A., Schein J.E.,  
RA Jones S.J., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human  
RT and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Kidney;  
RA Director MGC Project;  
RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.  
DR EMBL; BC082032; AAH82032.1; -;  
DR InterPro; IPR001526; LY6 UPAR.  
DR Pfam; PF00021; UPAR\_LY6; 1.  
KW Hypothetical protein.  
SQ SEQUENCE 141 AA; 15330 MW; BB8CC9993E322DA2 CRC64;  
Query Match 33.3%; Score 18; DB 2; Length 141;  
Best Local Similarity 28.6%; Pred. No. 1.5e+03;  
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
OY 3 CXXXXXX 9

DB 71 CASSAAC 77  
RESULT 22  
ID Q8BLC3 PRELIMINARY; PRT; 141 AA.  
AC Q8BLC3;  
DT 01-MAR-2003 (TREMBLrel. 23, Created)  
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)  
DT 25-OCT-2004 (TREMBLrel. 28, Last annotation update)  
DE Mus musculus adult male corpora quadrigemina cDNA, RIKEN full-length  
DE enriched library, clone:B230214B19 product:hypothetical Snake toxin-  
DE like structure containing protein, full insert sequence (Similar to  
DE human MGC39643 protein).  
GN Name=2700050C12Rik;  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Corpora quadrigemina;  
RX MEDLINE=9279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;  
RA Carninci P., Hayashizaki Y.;  
RT "High-efficiency full-length cDNA cloning.";  
RL Meth. Enzymol. 303:19-44(1999).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Corpora quadrigemina;  
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/350555500;  
RA RIKEN FANTOM Consortium;  
RT "Functional annotation of a full-length mouse cDNA collection.";  
RL Nature 409:685-690(2001).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Corpora quadrigemina;  
RX MEDLINE=20493374; PubMed=11042159; DOI=10.1101/gr.145100;  
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,  
RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;  
RT "Normalization and subtraction of cap-trapper-selected cDNAs to  
RT prepare full-length cDNA libraries for rapid discovery of new genes.";  
RL Genome Res. 10:1617-1630(2000).  
RN [5]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Corpora quadrigemina;  
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;  
RA Shibata K., Itoh M., Aizawa K., Nagao S., Sasaki N., Carninci P.,  
RA Konno H., Akiyama J., Nishi K., Kitauai T., Tashiro H., Itoh M.,  
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,  
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,  
RA Fujiwara S., Inoue K., Izawa M., Tanaka T., Matsura S., Watahiki M.,  
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Tanaka T., Matsura S., Watahiki M.,  
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;  
RT "RIKEN integrated sequence analysis (RISA) system-384-format  
RT sequencing pipeline with 384 multipipillary sequencer.";  
RL Genome Res. 10:1757-1771(2000).  
RN [6]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Corpora quadrigemina;  
RX Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,  
RA Fukuda S., Furuno M., Hanganaki T., Hara A., Hashizume W.,  
RA Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T.,  
RA Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,  
RA Katoh H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,

RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,  
RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,  
RA Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,  
RA Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,  
RA Tagawa A., Tanahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,  
RA Tomaru A., Toyota T., Yasunishi A., Muramatsu M., Hayaishizaki Y.,  
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.  
[7]  
RN  
RP SEQUENCE FROM N.A.  
RC TISSUE=Olfactory epithelium;  
RX MEDLINE=22389257; PubMed=12477932; DOI=10.1073/pnas.2426038999;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.J., Wang J., Hsieh P.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldi M.P., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalon D.K., Muzny D.N., Sodergren E.J., Lu X., Gibbs R.A., Sanchez A.,  
RA Fahey J., Helton E., Kettman M., Madan A.C., Rodrigues S.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
RA Krywinski M.I., Skalska J., Smailus D.E., Schnerch A., Schein J.E.,  
RA Jones S.J., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human  
and mouse cDNA sequences";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
[8]  
RN  
RP SEQUENCE FROM N.A.  
RC TISSUE=Olfactory epithelium;  
RA Strausberg R.;  
RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AK045592; BAC32428.1; -;  
DR EMBL; BC058599; AAH58599.1; -;  
DR MGD; MGI:1919835; 2700050C12Rik.  
DR GO; GO:0016526; F:G-protein coupled receptor activity, unknown. . .; TAS.  
DR GO; GO:0007186; P:G-protein coupled receptor protein signalin. . .; TAS.  
KW Hypothetical protein.  
SQ SEQUENCE 141 AA; 15261 MW; 51717F7922C02DAA CRC64;  
Query Match 33.3%; Score 18; DB 2; Length 141;  
Best Local Similarity 28.6%; Pred. No. 1.5e+03;  
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
Qy 3 CXXXXXC 9  
Db 71 CASSAAC 77  
RESULT 23  
Q9JJ96 PRELIMINARY; PRT; 141 AA.  
AC Q9JJ96  
DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Mus musculus brain cDNA, clone MNCB-0671.  
GN Name=2700050C12Rik;  
OS Mus musculus (Mouse)  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL;  
RA Osada N., Kusuda J., Tanuma R., Ito A., Hirata M., Sugano S.,  
RA Hashimoto K.;  
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AB041649; BAA95101.1; -;

DR MGD; MGI:1919835; 2700050C12Rik.  
DR GO; GO:0016526; F:G-protein coupled receptor activity, unknown. . .; TAS.  
DR GO; GO:0007186; P:G-protein coupled receptor protein signalin. . .; TAS.  
DR InterPro; IPR001526; LY6\_UPAR.  
DR SMART; SM00134; LU; 1.  
SQ SEQUENCE 141 AA; 15231 MW; C8716DDC652288ED CRC64;  
Query Match 33.3%; Score 18; DB 2; Length 141;  
Best Local Similarity 28.6%; Pred. No. 1.5e+03;  
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
Qy 3 CXXXXXC 9  
Db 71 CASSAAC 77  
RESULT 24  
UL42 HCMVA  
ID UL42 HCMVA STANDARD; PRT; 157 AA.  
AC P16815;  
DT 01-AUG-1990 (Rel. 15, Created)  
DT 01-AUG-1990 (Rel. 15, Last sequence update)  
DT 05-JUL-2004 (Rel. 44, Last annotation update)  
DE Hypothetical protein UL42.  
GN Name=UL42;  
OS Human cytomegalovirus (strain AD169).  
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;  
OC Betaherpesvirinae; Cytomegalovirus.  
OX NCBI\_TaxID=10360;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=90269039; PubMed=2161319;  
RA Chee M.S., Bankier A.T., Beck S., Bohni R., Brown C.M., Cerny R.,  
RA Horsnell T., Hutchison C.A. III, Kourazides T., Martignetti J.A.,  
RA Preddie E., Satchwell S.C., Tomlinson P., Weston K.M., Barrall B.G.;  
RT "Analysis of the protein-coding content of the sequence of human  
cytomegalovirus strain AD169";  
RL Curr. Top. Microbiol. Immunol. 154:125-169 (1990).  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; X17403; CAA35401.1; -;  
DR PIR; S09805; S09805.  
KW Hypothetical protein.  
FT CARBOHYD 147 147  
SQ SEQUENCE 157 AA; 17066 MW; 59BFBI5443124954 CRC64;  
N-linked (GlcNAc. . .) (Potential).  
Query Match 33.3%; Score 18; DB 1; Length 157;  
Best Local Similarity 28.6%; Pred. No. 1.6e+03;  
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
Qy 3 CXXXXXC 9  
Db 76 CAATSSC 82  
RESULT 25  
Q6ZP52 PRELIMINARY; PRT; 157 AA.  
AC Q6ZP52  
DT 05-JUL-2004 (TrEMBLrel. 27, Created)  
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)  
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)  
DE Hypothetical protein FLJ26495.  
OS Homo sapiens (Human)  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

```

OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RA Tashiro H., Yamazaki M., Watanabe K., Kumagai A., Itakura S.,
RA Fukuzumi Y., Fujimori Y., Komiyama M., Suzuki Y., Hata H.,
RA Nakagawa K., Mizuno S., Morinaga M., Kawamura M., Sugiyama T.,
RA Irie R., Otsuki T., Sato H., Nishikawa T., Sugiyama A., Kawakami B.,
RA Nagai K., Isogai T., Sugano S.;
RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK130005; BAC85273.1; -;
SQ SEQUENCE 157 AA; 16919 MW; 126998C032DB6AC5 CRC64;

Query Match 33.3%; Score 18; DB 2; Length 157;
Best Local Similarity 28.6%; Pred. No. 1.6e+03;
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 CXXXXXC 9
DB 87 CASSAAC 93

RESULT 26
Q7QF28 PRELIMINARY; PRT; 160 AA.
AC Q7QF28;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE AGCP13679 (Fragment).
GN Name=agCG54781; ORFNames=ENSANGG00000010518;
OS Anopheles gambiae str. PEST.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Anopheles.
OX NCBI_TaxID=180454;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PEST;
RA Anopheles Genome Sequencing Consortium;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAB0100846; EAA06475.1; -;
DR InterPro; IPR008597; Destabilase.
DR InterPro; IPR002048; EF-hand.
DR Pfam; PF05497; Destabilase; 1.
DR PROSITE; PS00018; EF_HAND; UNKNOWN_1.
FT NON_TER 1
SQ SEQUENCE 160 AA; 17368 MW; BEFA5PB917ACC080 CRC64;

Query Match 33.3%; Score 18; DB 2; Length 160;
Best Local Similarity 28.6%; Pred. No. 1.6e+03;
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 CXXXXXC 9
DB 45 CSTSTTC 51

RESULT 27
Q88EY6 PRELIMINARY; PRT; 162 AA.
AC Q88EY6;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN OrderedLocustNames=PP4314;
OS Pseudomonas putida (strain KT2440).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=160488;

SEQUENCE FROM N.A.
RX MEDLINE=22423060; PubMed=12534463;
RA Nelson K.E., Weinel C., Paulsen I.T., Dodson R.J., Hilbert H.,
RA Martins dos Santos V.A.P., Fouts D.E., Gill S.R., Pop M., Holmes M.,
RA Brinkac L.M., Beanan M.J., Deboy R.T., Daugherty S.C., Kolonay J.F.,
RA Madupu R., Nelson W.C., White O., Peterson J.D., Khouri H.M.,
RA Hance I., Chris Lee P., Holtzapple E.K., Scanlan D., Tran K.,
RA Moazzes A., Uterback T.R., Rizzo M., Lee K., Kosack D., Moestl D.,
RA Wedler H., Lauber J., Stjepandic D., Hoheisel J., Straetz M., Heim S.,
RA Kiewitz C., Eisen J.A., Timmis K.N., Duesterhoeft A., Tuemmler B.,
RA Fraser C.M.;
RT "Complete genome sequence and comparative analysis of the
RT metabolically versatile Pseudomonas putida KT2440."
RL Environ. Microbiol. 4:799-808(2002).
DR EMBL; AE016790; AAN69894.1; -;
DR TIGR; PA3114; -;
DR GO; GO:0016846; F:carbon-sulfur lyase activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR006913; GFA.
DR Pfam; PF04828; DUF636; 1.
KW Complete proteome: Hypothetical protein.
SQ SEQUENCE 162 AA; 18077 MW; 6PF514D988E63BD3 CRC64;

Query Match 33.3%; Score 18; DB 2; Length 162;
Best Local Similarity 28.6%; Pred. No. 1.6e+03;
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 CXXXXXC 9
DB 2 CTSSTAC 8

RESULT 28
Q82WI4 PRELIMINARY; PRT; 164 AA.
AC Q82WI4;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein FLJ41033.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Ishibashi T., Kanehori K., Yosida M., Watanabe S., Ishida S., Ono Y.,
RA Hotuta T., Hiraoka S., Murakawa K., Takiguchi S., Kusano J., Chiba Y.,
RA Watanabe M., Fujimori K., Tanai H., Ishida M., Yamashita H., Chiba Y.,
RA Sugiyama T., Irie R., Otsuki T., Sato H., Ota T., Wakamatsu A.,
RA Ishii S., Yamamoto J., Isono Y., Kawai-Hio Y., Saito K., Nishikawa T.,
RA Kimura K., Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K.,
RA Wagatsuma M., Takahashi-Fujii A., Oshima A., Sugiyama A., Kawakami B.,
RA Suzuki Y., Sugano S., Nagahari K., Masuho Y., Nagai K., Isogai T.;
RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK123029; BAC85518.1; -;
SQ SEQUENCE 164 AA; 17762 MW; EF4A4BCB69DA1252 CRC64;

Query Match 33.3%; Score 18; DB 2; Length 164;
Best Local Similarity 28.6%; Pred. No. 1.6e+03;
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 CXXXXXC 9
DB 94 CASSAAC 100

RESULT 29
Q8N2G4 PRELIMINARY; PRT; 165 AA.
ID Q8N2G4
AC Q8N2G4;

```

DT 01-OCT-2002 (TRENBLrel. 22, Created)  
DT 01-OCT-2002 (TRENBLrel. 22, Last sequence update)  
DE 01-MAR-2004 (TRENBLrel. 26, Last annotation update)  
DE Hypothetical protein PSEC0181.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RN SEQUENCE FROM N.A.  
RC TISSUE=Ovary;  
RA Ota T., Nishikawa T., Suzuki Y., Kawai-Hio Y., Hayashi K., Ishii S.,  
RA Saito K., Yamamoto J., Wakamatsu A., Nagai T., Nakamura Y.,  
RA Nagahari K., Sugano S., Isogai T.;  
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AK075487; BAC11647.1; -;  
DR InterPro; IPR001526; LY6\_UPAR.  
DR SMART; SM00134; LU; 1.  
SQ SEQUENCE 165 AA; 17900 MW; 2B0B8D68044CB999 CRC64;  
Query Match 33.3%; Score 18; DB 2; Length 165;  
Best Local Similarity 28.6%; Pred. No. 1.6e+03;  
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
OY 3 CXXXXXC 9  
DB 71 CASSAAC 77  
RESULT 30  
Q24969 PRELIMINARY; PRT; 167 AA.  
AC Q24969;  
DT 01-NOV-1996 (TRENBLrel. 01, Created)  
DT 01-DEC-2001 (TRENBLrel. 19, Last sequence update)  
DE 01-MAR-2004 (TRENBLrel. 26, Last annotation update)  
DE Giardia lamblia WB putative surface antigen (Fragment).  
OS Giardia lamblia (Giardia intestinalis).  
OC Eukaryota; Diplomonadida; Hexamitidae; Giardiinae; Giardia.  
OX NCBI\_TaxID=5741;  
RN [1]  
RN SEQUENCE FROM N.A.  
RC STRAIN=WB;  
RA Hilario E., Gogarten J.;  
RL Submitted (JUN-1994) to the EMBL/GenBank/DBJ databases.  
DR EMBL; U10907; AAA19317.1; -;  
DR InterPro; IPR006058; 2Fe2S\_fcd\_BS.  
DR InterPro; IPR009030; Grow fac. recept.  
DR PROSITE; PS00197; 2FE2S\_FERREDOXIN; UNKNOWN\_1.  
FT NON\_TER 1\_1  
SQ SEQUENCE 167 AA; 17671 MW; FDAB0A43EC276EDC CRC64;  
Query Match 33.3%; Score 18; DB 2; Length 167;  
Best Local Similarity 28.6%; Pred. No. 1.6e+03;  
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
OY 3 CXXXXXC 9  
DB 71 CASSAAC 77

Search completed: April 27, 2005, 16:41:38  
Job time : 177 secs

Result No.	Query			DB	ID	Description
	Score	Match	Length			
1	18	33.3	15	4	US-09-142-524D-141	Sequence 141, Appl
2	18	33.3	15	4	US-09-142-524D-142	Sequence 142, Appl
3	18	33.3	24	4	US-09-731-242A-3	Sequence 3, Appl
4	18	33.3	34	4	US-09-270-767-45088	Sequence 45088, A
5	18	33.3	34	4	US-09-270-767-55804	Sequence 55804, A
6	18	33.3	36	4	US-09-816-721-1	Sequence 1, Appl
7	18	33.3	48	5	PCF-US96-01720-8	Sequence 8, Appl
8	18	33.3	50	3	US-08-900-230-58	Sequence 58, Appl
9	18	33.3	70	4	US-09-621-976-6084	Sequence 6084, Ap
10	18	33.3	90	4	US-09-419-381-90	Sequence 90, Appl
11	18	33.3	125	4	US-09-252-991A-27726	Sequence 27726, A
12	18	33.3	127	3	US-08-467-023-189	Sequence 189, App
13	18	33.3	132	4	US-09-252-991A-22556	Sequence 22556, A
14	18	33.3	134	4	US-09-252-991A-25069	Sequence 25069, A
15	18	33.3	138	4	US-09-252-991A-16583	Sequence 16583, A
16	18	33.3	139	4	US-09-252-991A-29513	Sequence 29513, A
17	18	33.3	141	4	US-09-825-294-215	Sequence 215, App
18	18	33.3	141	4	US-09-970-966-215	Sequence 215, App
19	18	33.3	143	4	US-09-252-991A-17322	Sequence 17322, A
20	18	33.3	148	4	US-09-252-991A-17476	Sequence 17476, A
21	18	33.3	148	4	US-09-252-991A-23009	Sequence 23009, A
22	18	33.3	152	4	US-09-252-991A-24277	Sequence 24277, A
23	18	33.3	156	4	US-09-252-991A-31875	Sequence 31875, A
24	18	33.3	158	4	US-09-252-991A-30633	Sequence 30633, A
25	18	33.3	163	2	US-08-727-688-25	Sequence 25, Appl
26	18	33.3	163	4	US-09-252-991A-29098	Sequence 29098, A
27	18	33.3	166	4	US-09-252-991A-27896	Sequence 27896, A

101 17 31.5 17 3 US-08-450-653-9  
102 17 31.5 18 2 US-08-169-948B-48  
103 17 31.5 18 2 US-08-448-873-48  
104 17 31.5 20 2 US-08-637-759B-145  
105 17 31.5 20 3 US-08-871-355A-145  
106 17 31.5 20 3 US-08-476-509B-47  
107 17 31.5 20 3 US-08-467-023-54  
108 17 31.5 20 3 US-09-201-945-145  
109 17 31.5 20 4 US-09-148-545-227  
110 17 31.5 24 4 US-09-900-230-34  
111 17 31.5 24 4 US-09-643-657-41  
112 17 31.5 25 2 US-08-754-431A-11  
113 17 31.5 25 3 US-08-900-230-38  
114 17 31.5 29 3 US-09-136-769A-5  
115 17 31.5 29 3 US-09-136-769A-16  
116 17 31.5 31 2 US-08-169-948B-2  
117 17 31.5 31 2 US-08-448-873-2  
118 17 31.5 31 4 US-08-382-452D-2  
119 17 31.5 31 4 US-09-916-494A-2  
120 17 31.5 32 1 US-07-952-735A-3  
121 17 31.5 32 1 US-07-952-735A-4  
122 17 31.5 32 2 US-08-595-868C-41  
123 17 31.5 32 3 US-09-139-819A-41  
124 17 31.5 32 4 US-09-750-913-41  
125 17 31.5 32 6 5183802-5  
126 17 31.5 32 6 5183802-5  
127 17 31.5 33 1 US-07-776-272-9  
128 17 31.5 33 2 US-08-169-948B-6  
129 17 31.5 33 2 US-08-448-873-6  
130 17 31.5 33 3 US-08-382-452D-6  
131 17 31.5 33 4 US-08-507-362A-2  
132 17 31.5 34 4 US-09-916-494A-6  
133 17 31.5 34 1 US-07-952-735A-1  
134 17 31.5 34 1 US-07-952-735A-2  
135 17 31.5 36 4 US-09-166-966E-5  
136 17 31.5 38 3 US-08-900-230-55  
137 17 31.5 39 2 US-08-169-948B-4  
138 17 31.5 39 2 US-08-448-873-4  
139 17 31.5 39 3 US-08-382-452D-4  
140 17 31.5 39 4 US-09-916-494A-4  
141 17 31.5 44 1 US-08-676-169-6  
142 17 31.5 44 3 US-09-063-431A-6  
143 17 31.5 44 4 US-09-166-966E-9  
144 17 31.5 45 1 US-08-451-947-97  
145 17 31.5 45 1 US-08-451-947-98  
146 17 31.5 45 1 US-08-451-947-99  
147 17 31.5 45 1 US-08-451-947-100  
148 17 31.5 45 2 US-08-424-826A-11  
149 17 31.5 45 2 US-08-424-826A-12  
150 17 31.5 45 2 US-08-424-826A-97

## ALIGNMENTS

RESULT 1  
US-09-142-524D-141  
; Sequence 141, Application US/09142524D  
; Patent No. 6719976  
; GENERAL INFORMATION:  
; APPLICANT: Sone, Toshio  
; APPLICANT: Kume, Akinori  
; APPLICANT: Dairiki, Kazuo  
; APPLICANT: Iwama, Akiko  
; APPLICANT: Kino, Kohsuke  
; TITLE OF INVENTION: Peptide-Based Immunotherapeutic Agent for Treating Allergic Disease  
; FILE REFERENCE: SPO-103  
; CURRENT APPLICATION NUMBER: US/09/142,524D  
; PRIOR FILING DATE: 1998-09-09  
; PRIOR APPLICATION NUMBER: PCT/JP97/00740  
; PRIOR FILING DATE: 1997-03-10  
; NUMBER OF SEQ ID NOS: 174  
; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 141  
; LENGTH: 15  
; TYPE: PRT  
; ORGANISM: Cryptomeria japonica  
; FEATURE:  
; NAME/KEY: MISC FEATURE  
; LOCATION: (1)-(15)  
; OTHER INFORMATION: Cryj2 peptide, Figure 2, Row 58  
US-09-142-524D-141

Query Match 33.3%; Score 18; DB 4; Length 15;  
Best Local Similarity 28.6%; Pred. No. 1.6e+03;  
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 CXXXXXC 9  
DB 8 CTSASAC 14

## RESULT 2

US-09-142-524D-142  
; Sequence 142, Application US/09142524D  
; Patent No. 6719976  
; GENERAL INFORMATION:  
; APPLICANT: Sone, Toshio  
; APPLICANT: Kume, Akinori  
; APPLICANT: Dairiki, Kazuo  
; APPLICANT: Iwama, Akiko  
; APPLICANT: Kino, Kohsuke  
; TITLE OF INVENTION: Peptide-Based Immunotherapeutic Agent for Treating Allergic Disease  
; FILE REFERENCE: SPO-103  
; CURRENT APPLICATION NUMBER: US/09/142,524D  
; CURRENT FILING DATE: 1998-09-09  
; PRIOR APPLICATION NUMBER: PCT/JP97/00740  
; PRIOR FILING DATE: 1997-03-10  
; NUMBER OF SEQ ID NOS: 174  
; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 142  
; LENGTH: 15  
; TYPE: PRT  
; ORGANISM: Cryptomeria japonica  
; FEATURE:  
; NAME/KEY: MISC FEATURE  
; LOCATION: (1)-(15)  
; OTHER INFORMATION: Cryj2 peptide, Figure 2, Row 59  
US-09-142-524D-142

Query Match 33.3%; Score 18; DB 4; Length 15;  
Best Local Similarity 28.6%; Pred. No. 1.6e+03;  
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 CXXXXXC 9  
DB 3 CTSASAC 9

## RESULT 3

US-09-731-242A-3  
; Sequence 3, Application US/09731242A  
; Patent No. 6759243  
; GENERAL INFORMATION:  
; APPLICANT: Kuran, David  
; APPLICANT: Wittrup, K. Dane  
; APPLICANT: Holler, Phillip  
; TITLE OF INVENTION: HIGH AFFINITY TCR PROTEINS AND METHODS  
; FILE REFERENCE: 89-99  
; CURRENT APPLICATION NUMBER: US/09/731,242A  
; CURRENT FILING DATE: 2000-12-06  
; PRIOR APPLICATION NUMBER: US 60/169,179  
; PRIOR FILING DATE: 1999-12-06  
; PRIOR APPLICATION NUMBER: US 09/009,388  
; PRIOR FILING DATE: 1998-01-20  
; NUMBER OF SEQ ID NOS: 53

; SEQ ID NO 141  
; LENGTH: 15  
; TYPE: PRT  
; ORGANISM: Cryptomeria japonica  
; FEATURE:  
; NAME/KEY: MISC FEATURE  
; LOCATION: (1)-(15)  
; OTHER INFORMATION: Cryj2 peptide, Figure 2, Row 59  
US-09-142-524D-141

Query Match 33.3%; Score 18; DB 4; Length 15;  
Best Local Similarity 28.6%; Pred. No. 1.6e+03;  
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 CXXXXXC 9  
DB 3 CTSASAC 9

SOFTWARE: PatentIn version 3.0

SEQ ID NO 3

LENGTH: 24

TYPE: PRT

ORGANISM: ARTIFICIAL SEQUENCE

FEATURE:

NAME/KEY: misc\_feature

LOCATION: ( )..( )

OTHER INFORMATION: upstream primer

US-09-731-242A-3

Query Match 33.3%; Score 18; DB 4; Length 24;

Best Local Similarity 28.6%; Pred. No. 2e+03;

Matches 2; Conservative 0; Mismatches 5; Indels 5; Gaps 0;

QY 3 CXXXXXC 9

DB 9 CATAAAC 15

#### RESULT 4

US-09-270-767-40588

Sequence 40588, Application US/09270767

Patent No. 6703491

GENERAL INFORMATION:

APPLICANT: Homburger et al.

TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster

FILE REFERENCE: File Reference: 7326-094

CURRENT APPLICATION NUMBER: US/09/270,767

CURRENT FILING DATE: 1999-03-17

NUMBER OF SEQ ID NOS: 62517

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 40588

LENGTH: 34

TYPE: PRT

ORGANISM: Drosophila melanogaster

US-09-270-767-40588

Query Match 33.3%; Score 18; DB 4; Length 34;

Best Local Similarity 28.6%; Pred. No. 2.4e+03;

Matches 2; Conservative 0; Mismatches 5; Indels 5; Gaps 0;

QY 3 CXXXXXC 9

DB 4 CSSSSC 10

#### RESULT 5

US-09-270-767-55804

Sequence 55804, Application US/09270767

Patent No. 6703491

GENERAL INFORMATION:

APPLICANT: Homburger et al.

TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster

FILE REFERENCE: File Reference: 7326-094

CURRENT APPLICATION NUMBER: US/09/270,767

CURRENT FILING DATE: 1999-03-17

NUMBER OF SEQ ID NOS: 62517

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 55804

LENGTH: 34

TYPE: PRT

ORGANISM: Drosophila melanogaster

US-09-270-767-55804

Query Match 33.3%; Score 18; DB 4; Length 34;

Best Local Similarity 28.6%; Pred. No. 2.4e+03;

Matches 2; Conservative 0; Mismatches 5; Indels 5; Gaps 0;

QY 3 CXXXXXC 9

DB 4 CSSSSC 10

#### RESULT 6

US-09-816-721-1

Sequence 1, Application US/09816721

Patent No. 6777215

GENERAL INFORMATION:

APPLICANT: No. 6777215ozymes A/S

APPLICANT: Andersen, Carsten

APPLICANT: Nielsen, Bjarne R.

APPLICANT: Haren, Lubbet D.

APPLICANT: Dijkstra, Bauke H.

TITLE OF INVENTION: No. 6777215el Cyclomaltodextrin Glucanotransferase Variants

FILE REFERENCE: 5347.210-US

CURRENT APPLICATION NUMBER: US/09/816,721

CURRENT FILING DATE: 2001-03-22

NUMBER OF SEQ ID NOS: 2

SOFTWARE: PatentIn version 3.1

SEQ ID NO 1

LENGTH: 36

TYPE: PRT

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: synthetic

NAME/KEY: misc\_feature

OTHER INFORMATION: "DI96H oligo"

US-09-816-721-1

Query Match 33.3%; Score 18; DB 4; Length 36;

Best Local Similarity 28.6%; Pred. No. 2.5e+03;

Matches 2; Conservative 0; Mismatches 5; Indels 5; Gaps 0;

QY 3 CXXXXXC 9

DB 25 CTAAATC 31

#### RESULT 7

PCT-US96-01720-8

Sequence 8, Application PC/TUS9601720

GENERAL INFORMATION:

APPLICANT:

TITLE OF INVENTION: MODIFIED-AFFINITY STREPTAVIDIN

NUMBER OF SEQUENCES: 11

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US96/01720

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/387,055

FILING DATE: 09-FEB-1995

ATTORNEY/AGENT INFORMATION:

NAME: Parmelee, Steven W.

REGISTRATION NUMBER: 31,990

REFERENCE/DOCKET NUMBER: 16336-SPC

INFORMATION FOR SEQ ID NO: 8:

SEQUENCE CHARACTERISTICS:

LENGTH: 48 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

PCT-US96-01720-8

Query Match 33.3%; Score 18; DB 5; Length 48;

Best Local Similarity 28.6%; Pred. No. 2.9e+03;

Matches 2; Conservative 0; Mismatches 5; Indels 5; Gaps 0;

QY 3 CXXXXXC 9

```
Db      6 CTAATC 12

RESULT 8
US-08-900-230-58
; Sequence 58, Application US/08900230
; Patent No. 6329197
; GENERAL INFORMATION:
; APPLICANT: Bard, Jonathan A.
; TITLE OF INVENTION: DNA ENCODING GALANN GALR3 RECEPTORS AND
; TITLE OF INVENTION: USES THEREOF
; NUMBER OF SEQUENCES: 59
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 11036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/900,230
; FILING DATE: 23-JUL-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 52241-C/JPW/ADM
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-278-0400
; TELEFAX: 212-391-0525
; INFORMATION FOR SEQ ID NO: 58:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 50 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE:
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-08-900-230-58

Query Match      33.3%; Score 18; DB 3; Length 50;
Best Local Similarity 28.6%; Pred. No. 3e+03;
Matches 2; Conservative 0; Mismatches 5; Indels 5; Gaps 0;

QY      3 CXXXXXC 9
Db      9 CTAATAC 15

RESULT 9
US-09-621-976-6084
; Sequence 6084, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 6084
; LENGTH: 70
; TYPE: PRT

US-09-621-976-6084
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: 68
; OTHER INFORMATION: Xaa = Ala,Thr
; NAME/KEY: UNSURE
; LOCATION: 69
; OTHER INFORMATION: Xaa = Gly,Arg
; NAME/KEY: UNSURE
; LOCATION: 18
; OTHER INFORMATION: Xaa = Lys,Arg
US-09-621-976-6084

Query Match      33.3%; Score 18; DB 4; Length 70;
Best Local Similarity 28.6%; Pred. No. 3.5e+03;
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY      3 CXXXXXC 9
Db      42 CSASTSC 48

RESULT 10
US-09-419-381-90
; Sequence 90, Application US/09419381
; Patent No. 6747135
; GENERAL INFORMATION:
; APPLICANT: No. 6747135an, Garry P.
; APPLICANT: Rozinov, Michael N.
; TITLE OF INVENTION: Fluorescent Dye Binding Peptides
; FILE REFERENCE: A65681-1/DJB/RMS/DSS
; CURRENT APPLICATION NUMBER: US/09/419,381
; CURRENT FILING DATE: 1999-10-15
; PRIOR APPLICATION NUMBER: 60/104,465
; PRIOR FILING DATE: 1998-10-16
; NUMBER OF SEQ ID NOS: 122
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 90
; LENGTH: 90
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
US-09-419-381-90

Query Match      33.3%; Score 18; DB 4; Length 90;
Best Local Similarity 28.6%; Pred. No. 4e+03;
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY      3 CXXXXXC 9
Db      45 CAAATC 51

RESULT 11
US-09-252-991A-27726
; Sequence 27726, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 27726
; LENGTH: 125
; TYPE: PRT
```



ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-27726

Query Match 33.3%; Score 18; DB 4; Length 125;  
Best Local Similarity 28.6%; Pred. No. 4.7e+03;  
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 CXXXXXC 9  
DB 17 CTAATTC 23

## RESULT 12

US-08-467-023-189  
Sequence 189, Application US/08467023  
Patent No. 6090386

## GENERAL INFORMATION:

APPLICANT: Griffith, Irwin J.;  
APPLICANT: Pollock, Joanne;  
APPLICANT: Bond, Julian F.;  
APPLICANT: Garman, Richard D;  
APPLICANT: Kuo, Mei-Chang;  
APPLICANT: Yeung, Siu-mei H.;  
APPLICANT: Brauer, Andrew;  
APPLICANT: Exley, Mark A.;  
APPLICANT: Powers, Steven P.  
TITLE OF INVENTION: Allergenic Proteins And Peptides From  
TITLE OF INVENTION: Japanese Cedar Pollen  
NUMBER OF SEQUENCES: 261  
CORRESPONDENCE ADDRESS:  
ADDRESS: Immunologic Pharmaceutical Corporation, Inc.  
STREET: 610 Lincoln St  
CITY: Waltham  
STATE: MA  
COUNTRY: USA  
ZIP: 02154

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25

## CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/467,023  
FILING DATE: June 6, 1995  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/350,225  
FILING DATE: December 6, 1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Jane E. Remillard  
REGISTRATION NUMBER: 38,872  
REFERENCE/DOCKET NUMBER: 025.6 USD2 (IMI-028CPD2)  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 227-7400  
TELEFAX: (617) 227-5941

## INFORMATION FOR SEQ ID NO: 189:

SEQUENCE CHARACTERISTICS:  
LENGTH: 127 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FRAGMENT TYPE: internal  
US-08-467-023-189

Query Match 33.3%; Score 18; DB 3; Length 127;  
Best Local Similarity 28.6%; Pred. No. 4.8e+03;  
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 CXXXXXC 9  
DB 77 CTSASAC 83

## RESULT 13

US-09-252-991A-22556  
Sequence 22556, Application US/09252991A  
Patent No. 6551795

## GENERAL INFORMATION:

APPLICANT: Marc J. Rubenfield et al.  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
FILE REFERENCE: 107196.136  
CURRENT APPLICATION NUMBER: US/09/252,991A  
CURRENT FILING DATE: 1999-02-18  
PRIOR APPLICATION NUMBER: US 60/074,788  
PRIOR FILING DATE: 1998-02-18  
PRIOR APPLICATION NUMBER: US 60/094,190  
PRIOR FILING DATE: 1998-07-27  
NUMBER OF SEQ ID NOS: 33142  
SEQ ID NO 22556  
LENGTH: 132  
TYPE: PRT

## ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-22556

## Query Match

33.3%; Score 18; DB 4; Length 132;  
Best Local Similarity 28.6%; Pred. No. 4.9e+03;  
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 CXXXXXC 9  
DB 33 CSAASAC 39

## RESULT 14

US-09-252-991A-25069  
Sequence 25069, Application US/09252991A  
Patent No. 6551795

## GENERAL INFORMATION:

APPLICANT: Marc J. Rubenfield et al.  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
FILE REFERENCE: 107196.136  
CURRENT APPLICATION NUMBER: US/09/252,991A  
CURRENT FILING DATE: 1999-02-18  
PRIOR APPLICATION NUMBER: US 60/074,788  
PRIOR FILING DATE: 1998-02-18  
PRIOR APPLICATION NUMBER: US 60/094,190  
PRIOR FILING DATE: 1998-07-27  
NUMBER OF SEQ ID NOS: 33142  
SEQ ID NO 25069  
LENGTH: 134  
TYPE: PRT

## ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-25069

## Query Match

33.3%; Score 18; DB 4; Length 134;  
Best Local Similarity 28.6%; Pred. No. 4.9e+03;  
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 CXXXXXC 9  
DB 11 CSAATTC 17

## RESULT 15

US-09-252-991A-16583  
Sequence 16583, Application US/09252991A  
Patent No. 6551795

## GENERAL INFORMATION:

APPLICANT: Marc J. Rubenfield et al.  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
FILE REFERENCE: 107196.136  
CURRENT APPLICATION NUMBER: US/09/252,991A  
CURRENT FILING DATE: 1999-02-18

;; PRIOR APPLICATION NUMBER: US 60/074,788  
;; PRIOR FILING DATE: 1998-02-18  
;; PRIOR APPLICATION NUMBER: US 60/094,190  
;; PRIOR FILING DATE: 1998-07-27  
;; NUMBER OF SEQ ID NOS: 33142  
;; SEQ ID NO 16583  
;; LENGTH: 138  
;; TYPE: PRT  
;; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-16583

Query Match 33.3%; Score 18; DB 4; Length 138;  
Best Local Similarity 28.6%; Pred. No. 5e+03;  
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 CXXXXXC 9  
DB 53 CASTTAC 59

## RESULT 16

US-09-252-991A-29513  
;; Sequence 29513, Application US/09252991A  
;; Patent No. 6551795  
;; GENERAL INFORMATION:  
;; APPLICANT: Marc J. Rubenfield et al.  
;; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
;; FILE REFERENCE: 107196.136  
;; CURRENT APPLICATION NUMBER: US/09/252,991A  
;; PRIOR FILING DATE: 1999-02-18  
;; PRIOR APPLICATION NUMBER: US 60/074,788  
;; PRIOR FILING DATE: 1998-02-18  
;; PRIOR APPLICATION NUMBER: US 60/094,190  
;; PRIOR FILING DATE: 1998-07-27  
;; NUMBER OF SEQ ID NOS: 33142  
;; SEQ ID NO 29513  
;; LENGTH: 139  
;; TYPE: PRT  
;; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-29513

Query Match 33.3%; Score 18; DB 4; Length 139;  
Best Local Similarity 28.6%; Pred. No. 5e+03;  
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 CXXXXXC 9  
DB 48 CSTSTAC 54

## RESULT 17

US-09-825-294-215  
;; Sequence 215, Application US/09825294  
;; Patent No. 6710170  
;; GENERAL INFORMATION:  
;; APPLICANT: Xu, Jiangchun  
;; APPLICANT: Stolck, John A.  
;; APPLICANT: Algate, Paul A.  
;; APPLICANT: Fling, Steven P.  
;; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE  
;; FILE REFERENCE: 210121.484C5  
;; CURRENT APPLICATION NUMBER: US/09/825,294  
;; CURRENT FILING DATE: 2001-04-03  
;; NUMBER OF SEQ ID NOS: 215  
;; SOFTWARE: FastSeq for Windows Version 3.0  
;; SEQ ID NO 215  
;; LENGTH: 141  
;; TYPE: PRT  
;; ORGANISM: Homo sapiens  
US-09-825-294-215

Query Match 33.3%; Score 18; DB 4; Length 141;  
Best Local Similarity 28.6%; Pred. No. 5.1e+03;  
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 CXXXXXC 9  
DB 71 CASSAAC 77

## RESULT 18

US-09-970-966-215  
;; Sequence 215, Application US/09970966  
;; Patent No. 6720146  
;; GENERAL INFORMATION:  
;; APPLICANT: Stolck, John A.  
;; APPLICANT: Moesh, David Alan  
;; APPLICANT: Fling, Steven P.  
;; APPLICANT: Xu, Jiangchun  
;; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
;; FILE REFERENCE: 210121.484C6  
;; CURRENT APPLICATION NUMBER: US/09/970,966  
;; CURRENT FILING DATE: 2001-10-02  
;; NUMBER OF SEQ ID NOS: 215  
;; SOFTWARE: FastSeq for Windows Version 4.0  
;; SEQ ID NO 215  
;; LENGTH: 141  
;; TYPE: PRT  
;; ORGANISM: Homo sapiens  
US-09-970-966-215

Query Match 33.3%; Score 18; DB 4; Length 141;  
Best Local Similarity 28.6%; Pred. No. 5.1e+03;  
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 CXXXXXC 9  
DB 71 CASSAAC 77

## RESULT 19

US-09-252-991A-17322  
;; Sequence 17322, Application US/09252991A  
;; Patent No. 6551795  
;; GENERAL INFORMATION:  
;; APPLICANT: Marc J. Rubenfield et al.  
;; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
;; FILE REFERENCE: 107196.136  
;; CURRENT APPLICATION NUMBER: US/09/252,991A  
;; PRIOR FILING DATE: 1999-02-18  
;; PRIOR APPLICATION NUMBER: US 60/074,788  
;; PRIOR FILING DATE: 1998-02-18  
;; PRIOR APPLICATION NUMBER: US 60/094,190  
;; PRIOR FILING DATE: 1998-07-27  
;; NUMBER OF SEQ ID NOS: 33142  
;; SEQ ID NO 17322  
;; LENGTH: 143  
;; TYPE: PRT  
;; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-17322

Query Match 33.3%; Score 18; DB 4; Length 143;  
Best Local Similarity 28.6%; Pred. No. 5.1e+03;  
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 CXXXXXC 9  
DB 114 CASSTAC 120

## RESULT 20

US-09-252-991A-17476

; Sequence 17476, Application US/09252991A  
; Patent No. 6551795

; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 107196.136

; CURRENT APPLICATION NUMBER: US/09/252,991A

; CURRENT FILING DATE: 1999-02-18

; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/094,190

; PRIOR FILING DATE: 1998-07-27

; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 17476

; LENGTH: 148

; TYPE: PRT

; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-17476

Query Match 33.3%; Score 18; DB 4; Length 148;

Best Local Similarity 28.6%; Pred. No. 5.2e+03;

Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 CXXXXXC 9

DB 18 CTTSSC 24

RESULT 21

US-09-252-991A-23009

; Sequence 23009, Application US/09252991A

; Patent No. 6551795

; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 107196.136

; CURRENT APPLICATION NUMBER: US/09/252,991A

; CURRENT FILING DATE: 1999-02-18

; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/094,190

; PRIOR FILING DATE: 1998-07-27

; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 23009

; LENGTH: 148

; TYPE: PRT

; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-23009

Query Match 33.3%; Score 18; DB 4; Length 148;

Best Local Similarity 28.6%; Pred. No. 5.2e+03;

Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 CXXXXXC 9

DB 68 CSSSAC 74

RESULT 22

US-09-252-991A-24277

; Sequence 24277, Application US/09252991A

; Patent No. 6551795

; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 107196.136

; CURRENT APPLICATION NUMBER: US/09/252,991A

; CURRENT FILING DATE: 1999-02-18

; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/074,788

; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27

; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 24277

; LENGTH: 152

; TYPE: PRT

; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-24277

Query Match 33.3%; Score 18; DB 4; Length 152;

Best Local Similarity 28.6%; Pred. No. 5.3e+03;

Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 CXXXXXC 9

DB 81 CATSSTC 87

RESULT 23

US-09-252-991A-31875

; Sequence 31875, Application US/09252991A

; Patent No. 6551795

; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 107196.136

; CURRENT APPLICATION NUMBER: US/09/252,991A

; CURRENT FILING DATE: 1999-02-18

; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/094,190

; PRIOR FILING DATE: 1998-07-27

; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 31875

; LENGTH: 156

; TYPE: PRT

; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-31875

Query Match 33.3%; Score 18; DB 4; Length 156;

Best Local Similarity 28.6%; Pred. No. 5.3e+03;

Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 CXXXXXC 9

DB 26 CSTTASC 32

RESULT 24

US-09-252-991A-30633

; Sequence 30633, Application US/09252991A

; Patent No. 6551795

; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 107196.136

; CURRENT APPLICATION NUMBER: US/09/252,991A

; CURRENT FILING DATE: 1999-02-18

; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/074,788

; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/094,190

; PRIOR FILING DATE: 1998-07-27

; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 30633

; LENGTH: 158

; TYPE: PRT

; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-30633

Query Match 33.3%; Score 18; DB 4; Length 158;

Best Local Similarity 28.6%; Pred. No. 5.4e+03;

Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
QY 3 CXXXXXC 9  
Db 2 CSSATTC 8

RESULT 25  
US-08-727-688-25  
; Sequence 25, Application US/08727688  
; Patent No. 5919638  
; GENERAL INFORMATION:  
; APPLICANT: Russell, John C.  
; TITLE OF INVENTION: Reagents and Methods for Detecting Prostate Tumors  
; NUMBER OF SEQUENCES: 36  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Abbott Laboratories  
; STREET: 100 Abbott Park Road D377/AP6D  
; CITY: Abbott Park  
; STATE: IL  
; COUNTRY: USA  
; ZIP: 60064  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: Fast-Seq Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/727,688  
; FILING DATE:  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Porembski, Priscilla E.  
; REGISTRATION NUMBER: 33,207  
; REFERENCE/DOCKET NUMBER: 5967.US.01  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (847) 937-0378  
; TELEFAX: (847) 938-2623  
; INFORMATION FOR SEQ ID NO: 25:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 163 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: No. 5919638e  
US-08-727-688-25

Query Match 33.3%; Score 18; DB 2; Length 163;  
Best Local Similarity 28.6%; Pred. No. 5.4e+03;  
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 CXXXXXC 9  
Db 72 CSSSSC 78

RESULT 26  
US-09-252-991A-29098  
; Sequence 29098, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 29098  
; LENGTH: 163  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-29098

Query Match 33.3%; Score 18; DB 4; Length 163;  
Best Local Similarity 28.6%; Pred. No. 5.4e+03;  
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 CXXXXXC 9  
Db 67 CSSATAC 73

RESULT 27  
US-09-252-991A-27896  
; Sequence 27896, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 27896  
; LENGTH: 166  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-27896

Query Match 33.3%; Score 18; DB 4; Length 166;  
Best Local Similarity 28.6%; Pred. No. 5.5e+03;  
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 CXXXXXC 9  
Db 29 CATSTAC 35

RESULT 28  
US-09-252-991A-25719  
; Sequence 25719, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 25719  
; LENGTH: 174  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-25719

Query Match 33.3%; Score 18; DB 4; Length 174;  
Best Local Similarity 28.6%; Pred. No. 5.6e+03;

Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 CXXXXXC 9  
|  
Db 24 CTASSAC 30

## RESULT 29

US-09-252-991A-31386  
; Sequence 31386, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; PRIOR FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 31386  
; LENGTH: 178  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-31386

Query Match 33.3%; Score 18; DB 4; Length 178;  
Best Local Similarity 28.6%; Pred. No. 5.7e+03;  
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 CXXXXXC 9  
|  
Db 29 CATASTC 35

## RESULT 30

US-09-252-991A-17994  
; Sequence 17994, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; PRIOR FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 17994  
; LENGTH: 190  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-17994

Query Match 33.3%; Score 18; DB 4; Length 190;  
Best Local Similarity 28.6%; Pred. No. 5.9e+03;  
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 CXXXXXC 9  
|  
Db 114 CASTTAC 120

Search completed: April 27, 2005, 16:45:58  
Job time : 47 secs

*This Page Blank (uspio)*

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 27, 2005, 16:38:46 ; Search time 133 Seconds  
(without alignments)

77.568 Million cell updates/sec

Title: US-09-732-348B-34

Perfect score: 54

Sequence: 1 XXCXXXXXXCXXXXXXXXXXXXXXXXXXXXXXX 31

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1424015 seqs, 332791073 residues

Total number of hits satisfying chosen parameters: 1424015

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 150 summaries

Database : Published Applications AA:\*

1:	/cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2:	/cgn2_6/ptodata/2/pubpaa/PT_NEW_PUB.pep.*
3:	/cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4:	/cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5:	/cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6:	/cgn2_6/ptodata/2/pubpaa/PTUS_PUBCOMB.pep.*
7:	/cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8:	/cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9:	/cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
10:	/cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
11:	/cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12:	/cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
13:	/cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
14:	/cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
15:	/cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
16:	/cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep.*
17:	/cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
18:	/cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
19:	/cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
20:	/cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	19	35.2	103	10	US-09-820-843A-77
2	18	33.3	10	10	US-09-572-404B-2385
3	18	33.3	10	10	US-09-572-404B-2387
4	18	33.3	10	10	US-09-572-404B-2399
5	18	33.3	10	10	US-09-572-404B-2401
6	18	33.3	10	10	US-09-572-404B-2425
7	18	33.3	10	10	US-09-572-404B-2485
8	18	33.3	10	10	US-09-572-404B-2527
9	18	33.3	10	10	US-09-572-404B-2533
10	18	33.3	10	10	US-09-572-404B-2543
11	18	33.3	10	10	US-09-572-404B-2557
12	18	33.3	12	16	US-10-763-992-32
13	18	33.3	15	14	US-10-354-240-141
14	18	33.3	15	14	US-10-354-240-141

14	18	33.3	15	14	US-10-354-240-142	Sequence 142, App
15	18	33.3	18	14	US-10-225-567A-1778	Sequence 1778, App
16	18	33.3	24	9	US-09-731-242A-3	Sequence 3, Appl
17	18	33.3	25	15	US-10-361-811-292	Sequence 292, App
18	18	33.3	25	15	US-10-369-186-292	Sequence 292, App
19	18	33.3	28	9	US-09-864-761-41104	Sequence 41104, A
20	18	33.3	33	16	US-10-697-399-10	Sequence 10, Appl
21	18	33.3	36	9	US-09-816-721-1	Sequence 1, Appl
22	18	33.3	39	14	US-10-029-386-31509	Sequence 31509, A
23	18	33.3	50	8	US-08-900-230-58	Sequence 58, Appl
24	18	33.3	58	14	US-10-029-386-28972	Sequence 28972, A
25	18	33.3	66	16	US-10-437-963-193586	Sequence 193586, A
26	18	33.3	67	15	US-10-424-599-202235	Sequence 202235, A
27	18	33.3	70	15	US-10-424-599-261219	Sequence 261219, A
28	18	33.3	71	14	US-10-011-931-32	Sequence 32, Appl
29	18	33.3	71	14	US-10-269-806-45	Sequence 45, Appl
30	18	33.3	71	14	US-10-269-806-61	Sequence 61, Appl
31	18	33.3	71	14	US-10-269-806-65	Sequence 65, Appl
32	18	33.3	71	14	US-10-269-806-77	Sequence 77, Appl
33	18	33.3	71	14	US-10-269-806-85	Sequence 85, Appl
34	18	33.3	71	14	US-10-269-806-109	Sequence 109, App
35	18	33.3	71	14	US-10-269-806-117	Sequence 117, App
36	18	33.3	71	14	US-10-269-806-133	Sequence 133, App
37	18	33.3	75	9	US-09-758-140-11	Sequence 11, Appl
38	18	33.3	77	16	US-10-437-963-199736	Sequence 199736, A
39	18	33.3	81	15	US-10-131-487A-107	Sequence 107, App
40	18	33.3	85	15	US-10-424-599-150900	Sequence 150900, A
41	18	33.3	86	9	US-09-925-300-1139	Sequence 1139, App
42	18	33.3	86	15	US-10-424-599-251581	Sequence 251581, A
43	18	33.3	90	16	US-10-437-963-200032	Sequence 200032, A
44	18	33.3	90	16	US-10-692-151-90	Sequence 90, Appl
45	18	33.3	91	16	US-10-767-701-56213	Sequence 56213, A
46	18	33.3	104	11	US-09-864-408A-934	Sequence 934, App
47	18	33.3	104	15	US-10-424-599-205166	Sequence 205166, A
48	18	33.3	104	15	US-10-424-599-232629	Sequence 232629, A
49	18	33.3	106	16	US-10-767-701-51683	Sequence 51683, A
50	18	33.3	114	15	US-10-361-811-289	Sequence 289, App
51	18	33.3	115	15	US-10-361-811-290	Sequence 290, App
52	18	33.3	115	15	US-10-369-186-290	Sequence 290, App
53	18	33.3	120	15	US-10-424-599-208557	Sequence 208557, A
54	18	33.3	121	16	US-10-767-701-50733	Sequence 50733, A
55	18	33.3	121	17	US-10-491-853-23	Sequence 23, Appl
56	18	33.3	123	16	US-10-437-963-161257	Sequence 161257, A
57	18	33.3	131	15	US-10-108-260A-4406	Sequence 4406, App
58	18	33.3	135	9	US-09-929-629-6	Sequence 6, Appl
59	18	33.3	137	14	US-10-212-677-284	Sequence 284, App
60	18	33.3	137	15	US-10-361-811-284	Sequence 284, App
61	18	33.3	137	15	US-10-369-186-284	Sequence 284, App
62	18	33.3	141	9	US-09-825-294-215	Sequence 215, App
63	18	33.3	141	9	US-09-970-966-215	Sequence 215, App
64	18	33.3	141	14	US-10-237-884-16	Sequence 16, Appl
65	18	33.3	141	14	US-10-230-163-16	Sequence 16, Appl
66	18	33.3	141	14	US-10-230-338-16	Sequence 16, Appl
67	18	33.3	141	14	US-10-218-631-16	Sequence 16, Appl
68	18	33.3	141	14	US-10-230-414-16	Sequence 16, Appl
69	18	33.3	141	14	US-10-232-224-16	Sequence 16, Appl
70	18	33.3	141	14	US-10-216-159A-16	Sequence 16, Appl
71	18	33.3	141	14	US-10-218-849-16	Sequence 16, Appl
72	18	33.3	141	14	US-10-227-873-16	Sequence 16, Appl
73	18	33.3	141	14	US-10-227-883-16	Sequence 16, Appl
74	18	33.3	141	14	US-10-219-076-16	Sequence 16, Appl
75	18	33.3	141	14	US-10-230-434-16	Sequence 16, Appl
76	18	33.3	141	14	US-10-219-003-16	Sequence 16, Appl
77	18	33.3	141	14	US-10-219-075-16	Sequence 16, Appl
78	18	33.3	141	14	US-10-219-464-16	Sequence 16, Appl
79	18	33.3	141	14	US-10-219-466-16	Sequence 16, Appl
80	18	33.3	141	14	US-10-219-479-16	Sequence 16, Appl
81	18	33.3	141	14	US-10-219-481-16	Sequence 16, Appl
82	18	33.3	141	14	US-10-230-260-16	Sequence 16, Appl
83	18	33.3	141	14	US-10-232-231-16	Sequence 16, Appl
84	18	33.3	141	14	US-10-232-233-16	Sequence 16, Appl
85	18	33.3	141	14	US-10-216-165-16	Sequence 16, Appl
86	18	33.3	141	14	US-10-216-165-16	Sequence 16, Appl

```
87 18 33.3 141 14 US-10-218-956-16 Sequence 16, Appl
88 18 33.3 141 14 US-10-219-468-16 Sequence 16, Appl
89 18 33.3 141 14 US-10-219-478-16 Sequence 16, Appl
90 18 33.3 141 14 US-10-219-536-16 Sequence 16, Appl
91 18 33.3 141 14 US-10-233-205-16 Sequence 16, Appl
92 18 33.3 141 14 US-10-219-072-16 Sequence 16, Appl
93 18 33.3 141 14 US-10-219-470-16 Sequence 16, Appl
94 18 33.3 141 14 US-10-219-474-16 Sequence 16, Appl
95 18 33.3 141 14 US-10-219-524-16 Sequence 16, Appl
96 18 33.3 141 14 US-10-219-528-16 Sequence 16, Appl
97 18 33.3 141 14 US-10-227-880-16 Sequence 16, Appl
98 18 33.3 141 14 US-10-227-881-16 Sequence 16, Appl
99 18 33.3 141 14 US-10-227-882-16 Sequence 16, Appl
100 18 33.3 141 14 US-10-230-436-16 Sequence 16, Appl
101 18 33.3 141 14 US-10-232-223-16 Sequence 16, Appl
102 18 33.3 141 14 US-10-232-225-16 Sequence 16, Appl
103 18 33.3 141 14 US-10-232-227-16 Sequence 16, Appl
104 18 33.3 141 14 US-10-232-229-16 Sequence 16, Appl
105 18 33.3 141 14 US-10-232-234-16 Sequence 16, Appl
106 18 33.3 141 14 US-10-219-060-16 Sequence 16, Appl
107 18 33.3 141 14 US-10-216-160-16 Sequence 16, Appl
108 18 33.3 141 14 US-10-216-162-16 Sequence 16, Appl
109 18 33.3 141 14 US-10-216-164-16 Sequence 16, Appl
110 18 33.3 141 14 US-10-216-167-16 Sequence 16, Appl
111 18 33.3 141 14 US-10-216-168-16 Sequence 16, Appl
112 18 33.3 141 14 US-10-219-065-16 Sequence 16, Appl
113 18 33.3 141 14 US-10-219-071-16 Sequence 16, Appl
114 18 33.3 141 14 US-10-219-074-16 Sequence 16, Appl
115 18 33.3 141 14 US-10-219-077-16 Sequence 16, Appl
116 18 33.3 141 14 US-10-219-465-16 Sequence 16, Appl
117 18 33.3 141 14 US-10-219-467-16 Sequence 16, Appl
118 18 33.3 141 14 US-10-219-469-16 Sequence 16, Appl
119 18 33.3 141 14 US-10-219-471-16 Sequence 16, Appl
120 18 33.3 141 14 US-10-219-473-16 Sequence 16, Appl
121 18 33.3 141 14 US-10-219-476-16 Sequence 16, Appl
122 18 33.3 141 14 US-10-219-482-16 Sequence 16, Appl
123 18 33.3 141 14 US-10-227-874-16 Sequence 16, Appl
124 18 33.3 141 14 US-10-227-876-16 Sequence 16, Appl
125 18 33.3 141 14 US-10-227-878-16 Sequence 16, Appl
126 18 33.3 141 14 US-10-229-974-16 Sequence 16, Appl
127 18 33.3 141 14 US-10-230-024-16 Sequence 16, Appl
128 18 33.3 141 14 US-10-230-113-16 Sequence 16, Appl
129 18 33.3 141 14 US-10-230-183-16 Sequence 16, Appl
130 18 33.3 141 14 US-10-230-234-16 Sequence 16, Appl
131 18 33.3 141 14 US-10-230-306-16 Sequence 16, Appl
132 18 33.3 141 14 US-10-230-426-16 Sequence 16, Appl
133 18 33.3 141 14 US-10-230-427-16 Sequence 16, Appl
134 18 33.3 141 14 US-10-230-433-16 Sequence 16, Appl
135 18 33.3 141 14 US-10-230-435-16 Sequence 16, Appl
136 18 33.3 141 14 US-10-230-438-16 Sequence 16, Appl
137 18 33.3 141 14 US-10-232-222-16 Sequence 16, Appl
138 18 33.3 141 14 US-10-219-070-16 Sequence 16, Appl
139 18 33.3 141 14 US-10-219-472-16 Sequence 16, Appl
140 18 33.3 141 14 US-10-219-527-16 Sequence 16, Appl
141 18 33.3 141 14 US-10-227-877-16 Sequence 16, Appl
142 18 33.3 141 14 US-10-216-166-16 Sequence 16, Appl
143 18 33.3 141 14 US-10-218-612-16 Sequence 16, Appl
144 18 33.3 141 14 US-10-212-677-215 Sequence 215, App
145 18 33.3 141 14 US-10-241-220-64 Sequence 64, Appl
146 18 33.3 141 14 US-10-216-163-16 Sequence 16, Appl
147 18 33.3 141 14 US-10-218-765-16 Sequence 16, Appl
148 18 33.3 141 14 US-10-219-063-16 Sequence 16, Appl
149 18 33.3 141 14 US-10-219-066-16 Sequence 16, Appl
150 18 33.3 141 14 US-10-219-067-16 Sequence 16, Appl
```

ALIGNMENTS

RESULT 1  
US-09-820-843A-77  
; Sequence 77, Application US/09820843A  
; Publication No. US200303963A1

GENERAL INFORMATION:  
; APPLICANT: Council of Scientific and Industrial Research  
; TITLE OF INVENTION: A COMPUTATIONAL METHOD FOR THE IDENTIFICATION OF CANDIDATE PROTEIN  
; FILE REFERENCE: Q63915  
; CURRENT APPLICATION NUMBER: US/09/820,843A  
; CURRENT FILING DATE: 2001-03-30  
; NUMBER OF SEQ ID NOS: 118  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 77  
; LENGTH: 103  
; TYPE: PRT  
; ORGANISM: M. pneumoniae  
; FEATURE:  
; NAME/KEY: misc feature  
; OTHER INFORMATION: Bol\_orf103b Protein  
; NAME/KEY: misc feature  
; OTHER INFORMATION: gi|1673772  
US-09-820-843A-77  
Query Match 35.2%; Score 19; DB 10; Length 103;  
Best Local Similarity 13.6%; Pred. No. 4.9e+03;  
Matches 3; Conservative 0; Mismatches 19; Indels 0; Gaps 0;  
QY 3 CXXXXXXCXXXXXXXXXXXXX 24  
DB 53 CSSISFCSSASSARLYSSSH 74.  
RESULT 2  
US-09-572-404B-2385  
; Sequence 2385, Application US/09572404B  
; Publication No. US20030078374A1  
; GENERAL INFORMATION:  
; APPLICANT: Proteom Ltd  
; TITLE OF INVENTION: Complementary peptide ligands from the human genome  
; FILE REFERENCE: Human patent  
; CURRENT APPLICATION NUMBER: US/09/572,404B  
; CURRENT FILING DATE: 2000-05-17  
; NUMBER OF SEQ ID NOS: 4203  
; SOFTWARE: ProtPatent version 1.0  
; SEQ ID NO 2385  
; LENGTH: 10  
; TYPE: PRT  
; ORGANISM: Homo Sapiens  
; FEATURE:  
; OTHER INFORMATION: sequence located in CBA at 82-91 and may interact with Sequence 2  
US-09-572-404B-2385  
Query Match 33.3%; Score 18; DB 10; Length 10;  
Best Local Similarity 28.6%; Pred. No. 2.4e+03;  
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
QY 3 CXXXXXC 9  
DB 3 CSSSTTC 9  
RESULT 3  
US-09-572-404B-2387  
; Sequence 2387, Application US/09572404B  
; Publication No. US20030078374A1  
; GENERAL INFORMATION:  
; APPLICANT: Proteom Ltd  
; TITLE OF INVENTION: Complementary peptide ligands from the human genome  
; FILE REFERENCE: Human patent  
; CURRENT APPLICATION NUMBER: US/09/572,404B  
; CURRENT FILING DATE: 2000-05-17  
; NUMBER OF SEQ ID NOS: 4203  
; SOFTWARE: ProtPatent version 1.0  
; SEQ ID NO 2387  
; LENGTH: 10



```
; TYPE: PRT
; ORGANISM: Homo Sapiens
; FEATURE:
; OTHER INFORMATION: sequence located in C8A at 81-90 and may interact with Sequence 2
; OTHER INFORMATION: this patent.
US-09-572-404B-2387

Query Match      33.3%; Score 18; DB 10; Length 10;
Best Local Similarity 28.6%; Pred. No. 2.4e+03;
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 CXXXXXC 9
Db 4 CSSSTTC 10

RESULT 4
US-09-572-404B-2399
; Sequence 2399, Application US/09572404B
; Publication No. US20030078374A1
; GENERAL INFORMATION:
; APPLICANT: Proteom Ltd
; TITLE OF INVENTION: Complementary peptide ligands from the human genome
; FILE REFERENCE: Human patent
; CURRENT APPLICATION NUMBER: US/09/572,404B
; CURRENT FILING DATE: 2000-05-17
; NUMBER OF SEQ ID NOS: 4203
; SOFTWARE: ProtPatent version 1.0
; SEQ ID NO 2399
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Homo Sapiens
; FEATURE:
; OTHER INFORMATION: sequence located in C8A at 81-90 and may interact with Sequence 2
; OTHER INFORMATION: this patent.
US-09-572-404B-2399

Query Match      33.3%; Score 18; DB 10; Length 10;
Best Local Similarity 28.6%; Pred. No. 2.4e+03;
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 CXXXXXC 9
Db 4 CSSSTTC 10

RESULT 5
US-09-572-404B-2401
; Sequence 2401, Application US/09572404B
; Publication No. US20030078374A1
; GENERAL INFORMATION:
; APPLICANT: Proteom Ltd
; TITLE OF INVENTION: Complementary peptide ligands from the human genome
; FILE REFERENCE: Human patent
; CURRENT APPLICATION NUMBER: US/09/572,404B
; CURRENT FILING DATE: 2000-05-17
; NUMBER OF SEQ ID NOS: 4203
; SOFTWARE: ProtPatent version 1.0
; SEQ ID NO 2401
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Homo Sapiens
; FEATURE:
; OTHER INFORMATION: sequence located in C8A at 83-92 and may interact with Sequence 2
; OTHER INFORMATION: this patent.
US-09-572-404B-2401

Query Match      33.3%; Score 18; DB 10; Length 10;
Best Local Similarity 28.6%; Pred. No. 2.4e+03;
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 CXXXXXC 9
Db 4 CSSSTTC 10

RESULT 6
US-09-572-404B-2425
; Sequence 2425, Application US/09572404B
; Publication No. US20030078374A1
; GENERAL INFORMATION:
; APPLICANT: Proteom Ltd
; TITLE OF INVENTION: Complementary peptide ligands from the human genome
; FILE REFERENCE: Human patent
; CURRENT APPLICATION NUMBER: US/09/572,404B
; CURRENT FILING DATE: 2000-05-17
; NUMBER OF SEQ ID NOS: 4203
; SOFTWARE: ProtPatent version 1.0
; SEQ ID NO 2425
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Homo Sapiens
; FEATURE:
; OTHER INFORMATION: sequence located in C8A at 84-93 and may interact with Sequence 2
; OTHER INFORMATION: this patent.
US-09-572-404B-2425

Query Match      33.3%; Score 18; DB 10; Length 10;
Best Local Similarity 28.6%; Pred. No. 2.4e+03;
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 CXXXXXC 9
Db 4 CSSSTTC 10

RESULT 7
US-09-572-404B-2485
; Sequence 2485, Application US/09572404B
; Publication No. US20030078374A1
; GENERAL INFORMATION:
; APPLICANT: Proteom Ltd
; TITLE OF INVENTION: Complementary peptide ligands from the human genome
; FILE REFERENCE: Human patent
; CURRENT APPLICATION NUMBER: US/09/572,404B
; CURRENT FILING DATE: 2000-05-17
; NUMBER OF SEQ ID NOS: 4203
; SOFTWARE: ProtPatent version 1.0
; SEQ ID NO 2485
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Homo Sapiens
; FEATURE:
; OTHER INFORMATION: sequence located in C8A at 81-90 and may interact with Sequence 2
; OTHER INFORMATION: this patent.
US-09-572-404B-2485

Query Match      33.3%; Score 18; DB 10; Length 10;
Best Local Similarity 28.6%; Pred. No. 2.4e+03;
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 CXXXXXC 9
Db 4 CSSSTTC 10

RESULT 8
US-09-572-404B-2527
; Sequence 2527, Application US/09572404B
; Publication No. US20030078374A1
; GENERAL INFORMATION:
; APPLICANT: Proteom Ltd
; TITLE OF INVENTION: Complementary peptide ligands from the human genome
; FILE REFERENCE: Human patent
; CURRENT APPLICATION NUMBER: US/09/572,404B
; CURRENT FILING DATE: 2000-05-17
; SEQ ID NO 2527
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Homo Sapiens
; FEATURE:
; OTHER INFORMATION: sequence located in C8A at 81-90 and may interact with Sequence 2
; OTHER INFORMATION: this patent.
US-09-572-404B-2527

Query Match      33.3%; Score 18; DB 10; Length 10;
Best Local Similarity 28.6%; Pred. No. 2.4e+03;
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 CXXXXXC 9
Db 4 CSSSTTC 10
```

```
Db 2 CSSSTTC 8

RESULT 6
US-09-572-404B-2425
; Sequence 2425, Application US/09572404B
; Publication No. US20030078374A1
; GENERAL INFORMATION:
; APPLICANT: Proteom Ltd
; TITLE OF INVENTION: Complementary peptide ligands from the human genome
; FILE REFERENCE: Human patent
; CURRENT APPLICATION NUMBER: US/09/572,404B
; CURRENT FILING DATE: 2000-05-17
; NUMBER OF SEQ ID NOS: 4203
; SOFTWARE: ProtPatent version 1.0
; SEQ ID NO 2425
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Homo Sapiens
; FEATURE:
; OTHER INFORMATION: sequence located in C8A at 84-93 and may interact with Sequence 2
; OTHER INFORMATION: this patent.
US-09-572-404B-2425

Query Match      33.3%; Score 18; DB 10; Length 10;
Best Local Similarity 28.6%; Pred. No. 2.4e+03;
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 CXXXXXC 9
Db 1 CSSSTTC 7

RESULT 7
US-09-572-404B-2485
; Sequence 2485, Application US/09572404B
; Publication No. US20030078374A1
; GENERAL INFORMATION:
; APPLICANT: Proteom Ltd
; TITLE OF INVENTION: Complementary peptide ligands from the human genome
; FILE REFERENCE: Human patent
; CURRENT APPLICATION NUMBER: US/09/572,404B
; CURRENT FILING DATE: 2000-05-17
; NUMBER OF SEQ ID NOS: 4203
; SOFTWARE: ProtPatent version 1.0
; SEQ ID NO 2485
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Homo Sapiens
; FEATURE:
; OTHER INFORMATION: sequence located in C8A at 81-90 and may interact with Sequence 2
; OTHER INFORMATION: this patent.
US-09-572-404B-2485

Query Match      33.3%; Score 18; DB 10; Length 10;
Best Local Similarity 28.6%; Pred. No. 2.4e+03;
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 CXXXXXC 9
Db 4 CSSSTTC 10

RESULT 8
US-09-572-404B-2527
; Sequence 2527, Application US/09572404B
; Publication No. US20030078374A1
; GENERAL INFORMATION:
; APPLICANT: Proteom Ltd
; TITLE OF INVENTION: Complementary peptide ligands from the human genome
; FILE REFERENCE: Human patent
; CURRENT APPLICATION NUMBER: US/09/572,404B
; CURRENT FILING DATE: 2000-05-17
; SEQ ID NO 2527
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Homo Sapiens
; FEATURE:
; OTHER INFORMATION: sequence located in C8A at 81-90 and may interact with Sequence 2
; OTHER INFORMATION: this patent.
US-09-572-404B-2527

Query Match      33.3%; Score 18; DB 10; Length 10;
Best Local Similarity 28.6%; Pred. No. 2.4e+03;
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 CXXXXXC 9
Db 4 CSSSTTC 10
```

```
/ NUMBER OF SEQ ID NOS: 4203
/ SOFTWARE: ProtPatent version 1.0
/ SEQ ID NO 2527
/ LENGTH: 10
/ TYPE: PRT
/ ORGANISM: Homo Sapiens
/ FEATURE:
/ OTHER INFORMATION: sequence located in C8A at 83-92 and may interact with Sequence 2
/ OTHER INFORMATION: this patent.
US-09-572-404B-2527

Query Match      33.3%; Score 18; DB 10; Length 10;
Best Local Similarity 28.6%; Pred. No. 2.4e+03;
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY      3 CXXXXXC 9
DB      2 CSSSTTC 8

RESULT 9
US-09-572-404B-2533
/ Sequence 2533, Application US/09572404B
/ Publication No. US20030078374A1
/ GENERAL INFORMATION:
/ APPLICANT: Proteom Ltd
/ TITLE OF INVENTION: Complementary peptide ligands from the human genome
/ FILE REFERENCE: Human patent
/ CURRENT APPLICATION NUMBER: US/09/572,404B
/ CURRENT FILING DATE: 2000-05-17
/ NUMBER OF SEQ ID NOS: 4203
/ SOFTWARE: ProtPatent version 1.0
/ SEQ ID NO 2533
/ LENGTH: 10
/ TYPE: PRT
/ ORGANISM: Homo Sapiens
/ FEATURE:
/ OTHER INFORMATION: sequence located in C8A at 83-92 and may interact with Sequence 2
/ OTHER INFORMATION: this patent.
US-09-572-404B-2533

Query Match      33.3%; Score 18; DB 10; Length 10;
Best Local Similarity 28.6%; Pred. No. 2.4e+03;
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY      3 CXXXXXC 9
DB      2 CSSSTTC 8

RESULT 10
US-09-572-404B-2543
/ Sequence 2543, Application US/09572404B
/ Publication No. US20030078374A1
/ GENERAL INFORMATION:
/ APPLICANT: Proteom Ltd
/ TITLE OF INVENTION: Complementary peptide ligands from the human genome
/ FILE REFERENCE: Human patent
/ CURRENT APPLICATION NUMBER: US/09/572,404B
/ CURRENT FILING DATE: 2000-05-17
/ NUMBER OF SEQ ID NOS: 4203
/ SOFTWARE: ProtPatent version 1.0
/ SEQ ID NO 2543
/ LENGTH: 10
/ TYPE: PRT
/ ORGANISM: Homo Sapiens
/ FEATURE:
/ OTHER INFORMATION: sequence located in C8A at 81-90 and may interact with Sequence 2
/ OTHER INFORMATION: this patent.
US-09-572-404B-2543

Query Match      33.3%; Score 18; DB 10; Length 10;
Best Local Similarity 28.6%; Pred. No. 2.4e+03;
```

```
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY      3 CXXXXXC 9
DB      4 CSSSTTC 10

RESULT 11
US-09-572-404B-2557
/ Sequence 2557, Application US/09572404B
/ Publication No. US20030078374A1
/ GENERAL INFORMATION:
/ APPLICANT: Proteom Ltd
/ TITLE OF INVENTION: Complementary peptide ligands from the human genome
/ FILE REFERENCE: Human patent
/ CURRENT APPLICATION NUMBER: US/09/572,404B
/ CURRENT FILING DATE: 2000-05-17
/ NUMBER OF SEQ ID NOS: 4203
/ SOFTWARE: ProtPatent version 1.0
/ SEQ ID NO 2557
/ LENGTH: 10
/ TYPE: PRT
/ ORGANISM: Homo Sapiens
/ FEATURE:
/ OTHER INFORMATION: sequence located in C8A at 83-92 and may interact with Sequence 2
/ OTHER INFORMATION: this patent.
US-09-572-404B-2557

Query Match      33.3%; Score 18; DB 10; Length 10;
Best Local Similarity 28.6%; Pred. No. 2.4e+03;
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY      3 CXXXXXC 9
DB      2 CSSSTTC 8

RESULT 12
US-10-763-992-32
/ Sequence 32, Application US/10763992
/ Publication No. US20040121397A1
/ GENERAL INFORMATION:
/ APPLICANT: COHEN, Maurice
/ FRIEDMAN, Paula N.
/ GORDON, Julian
/ HODGES, Steven C.
/ KLASS, Michael R.
/ KRATOCHVIL, Jon D.
/ ROBERTS-RAPP, Lisa
/ RUSSELL, John C.
/ STROUPE, Steven D.
/ TITLE OF INVENTION: REAGENTS AND METHODS USEFUL
/ FOR DETECTING DISEASES OF THE PROSTATE
/ NUMBER OF SEQUENCES: 35
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Abbott Laboratories
/ STREET: 100 Abbott Park Road
/ CITY: Abbott Park
/ STATE: IL
/ COUNTRY: USA
/ ZIP: 60064-3500
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Diskette
/ COMPUTER: IBM Compatible
/ OPERATING SYSTEM: DOS
/ SOFTWARE: FastSeq for Windows Version 2.0
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/10/763,992
/ FILING DATE: 22-Jan-2004
/ CLASSIFICATION: <Unknown>
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US/09/418,887
/ FILING DATE: 15-OCT-1999
```

APPLICATION NUMBER: US/08/946,869  
FILING DATE: 08-Oct-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Becker, Cheryl L.  
REGISTRATION NUMBER: 35,441  
REFERENCE/DOCKET NUMBER: 5697.US.P1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 847/935-1729  
TELEFAX: 847/938-2623  
TELEX: <Unknown>  
INFORMATION FOR SEQ ID NO: 32:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 12 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: None  
SEQUENCE DESCRIPTION: SEQ ID NO: 32:  
US-10-763-992-32

Query Match 33.3%; Score 18; DB 16; Length 12;  
Best Local Similarity 28.6%; Pred. No. 2.7e+03;  
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 CXXXXXC 9  
DB 1 CSSSSC 7

RESULT 13  
US-10-354-240-141  
Sequence 141, Application US/10354240  
Publication No. US20030185847A1  
GENERAL INFORMATION:

APPLICANT: Sone, Toshio  
APPLICANT: Kume, Akinori  
APPLICANT: Dairiki, Kazuo  
APPLICANT: Iwama, Akiko  
APPLICANT: Kino, Kohsuke  
TITLE OF INVENTION: Peptide-Based Immunotherapeutic Agent for Treating Allergic Disease  
FILE REFERENCE: SPO-103D1  
CURRENT APPLICATION NUMBER: US/10/354,240  
CURRENT FILING DATE: 2003-01-29  
PRIOR APPLICATION NUMBER: PCT/JP97/00740  
PRIOR FILING DATE: 1997-03-10  
PRIOR APPLICATION NUMBER: US 09/142,524  
PRIOR FILING DATE: 1998-09-09  
NUMBER OF SEQ ID NOS: 174  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 141  
LENGTH: 15  
TYPE: PRT  
ORGANISM: Cryptomeria japonica  
FEATURE:  
NAME/KEY: MISC FEATURE  
LOCATION: (1)..(15)  
OTHER INFORMATION: Cryj2 peptide, Figure 2, Row 58  
US-10-354-240-141

Query Match 33.3%; Score 18; DB 14; Length 15;  
Best Local Similarity 28.6%; Pred. No. 3e+03;  
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 CXXXXXC 9  
DB 8 CTSASAC 14

RESULT 14  
US-10-354-240-142  
Sequence 142, Application US/10354240  
Publication No. US20030185847A1  
GENERAL INFORMATION:

APPLICANT: Sone, Toshio  
APPLICANT: Kume, Akinori  
APPLICANT: Dairiki, Kazuo  
APPLICANT: Iwama, Akiko  
APPLICANT: Kino, Kohsuke  
TITLE OF INVENTION: Peptide-Based Immunotherapeutic Agent for Treating Allergic Disease  
FILE REFERENCE: SPO-103D1  
CURRENT APPLICATION NUMBER: US/10/354,240  
CURRENT FILING DATE: 2003-01-29  
PRIOR APPLICATION NUMBER: PCT/JP97/00740  
PRIOR FILING DATE: 1997-03-10  
PRIOR APPLICATION NUMBER: US 09/142,524  
PRIOR FILING DATE: 1998-09-09  
NUMBER OF SEQ ID NOS: 174  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 142  
LENGTH: 15  
TYPE: PRT  
ORGANISM: Cryptomeria japonica  
FEATURE:  
NAME/KEY: MISC FEATURE  
LOCATION: (1)..(15)  
OTHER INFORMATION: Cryj2 peptide, Figure 2, Row 59  
US-10-354-240-142

Query Match 33.3%; Score 18; DB 14; Length 15;  
Best Local Similarity 28.6%; Pred. No. 3e+03;  
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 CXXXXXC 9  
DB 3 CTSASAC 9

RESULT 15  
US-10-225-567A-1778  
Sequence 1778, Application US/10225567A  
Publication No. US20030113798A1  
GENERAL INFORMATION:  
APPLICANT: LifeSpan Biosciences  
APPLICANT: Brown, Joseph P.  
APPLICANT: Burner, Glenna C.  
APPLICANT: Roush, Christine L.  
TITLE OF INVENTION: ANTIGENIC PEPTIDES AND ANTIBODIES FOR G PROTEIN-COUPLED RECEPTORS  
FILE REFERENCE: 1920-4-4  
CURRENT APPLICATION NUMBER: US/10/225,567A  
CURRENT FILING DATE: 2001-12-19  
PRIOR APPLICATION NUMBER: 60/257,144  
PRIOR FILING DATE: 2000-12-19  
NUMBER OF SEQ ID NOS: 2292  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 1778  
LENGTH: 18  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-225-567A-1778

Query Match 33.3%; Score 18; DB 14; Length 18;  
Best Local Similarity 28.6%; Pred. No. 3.3e+03;  
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 CXXXXXC 9  
DB 9 CASSSSC 15

RESULT 16  
US-09-731-242A-3  
Sequence 3, Application US/09731242A  
Patent No. US20020058253A1  
GENERAL INFORMATION:  
APPLICANT: KRANZ, DAVID  
APPLICANT: WITTRUP, K. DANE

APPLICANT: HOLLER, PHILLIP  
TITLE OF INVENTION: HIGH AFFINITY TCR PROTEINS AND METHODS  
FILE REFERENCE: 89-99  
CURRENT APPLICATION NUMBER: US/09/731,242A  
CURRENT FILING DATE: 2000-12-06  
PRIOR APPLICATION NUMBER: US 60/169,179  
PRIOR FILING DATE: 1999-12-06  
PRIOR APPLICATION NUMBER: US 09/009,388  
PRIOR FILING DATE: 1998-01-20  
NUMBER OF SEQ ID NOS: 53  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 3  
LENGTH: 24  
TYPE: PRT  
ORGANISM: ARTIFICIAL SEQUENCE  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: ()..()  
OTHER INFORMATION: upstream primer  
US-09-731-242A-3

Query Match 33.3%; Score 18; DB 9; Length 24;  
Best Local Similarity 28.6%; Pred. No. 3.8e+03;  
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 3 CXXXXXC 9  
Db 9 CATAAAC 15

RESULT 17  
US-10-361-811-292  
Sequence 292, Application US/10361811  
Publication No. US20030206918A1  
GENERAL INFORMATION:  
APPLICANT: Fanger, Gary R.  
APPLICANT: Fling, Steven P.  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
TITLE OF INVENTION: AND DIAGNOSIS OF OVARIAN CANCER  
FILE REFERENCE: 210121.484C8  
CURRENT APPLICATION NUMBER: US/10/361,811  
CURRENT FILING DATE: 2003-02-05  
NUMBER OF SEQ ID NOS: 293  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 292  
LENGTH: 25  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Synthetic peptide used for generation of rabbit  
OTHER INFORMATION: polyclonal anti-sera against O591a.  
US-10-361-811-292

Query Match 33.3%; Score 18; DB 15; Length 25;  
Best Local Similarity 28.6%; Pred. No. 3.9e+03;  
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 3 CXXXXXC 9  
Db 18 CASSAAC 24

RESULT 18  
US-10-369-186-292  
Sequence 292, Application US/10369186  
Publication No. US20030232056A1  
GENERAL INFORMATION:  
APPLICANT: Fanger, Gary R.  
APPLICANT: Fling, Steven P.  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
TITLE OF INVENTION: AND DIAGNOSIS OF OVARIAN CANCER  
FILE REFERENCE: 210121.484C9  
CURRENT APPLICATION NUMBER: US/10/369,186

CURRENT FILING DATE: 2003-02-14  
NUMBER OF SEQ ID NOS: 293  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 292  
LENGTH: 25  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Synthetic peptide used for generation of rabbit  
OTHER INFORMATION: polyclonal anti-sera against O591a.  
US-10-369-186-292

Query Match 33.3%; Score 18; DB 15; Length 25;  
Best Local Similarity 28.6%; Pred. No. 3.9e+03;  
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 3 CXXXXXC 9  
Db 18 CASSAAC 24

RESULT 19  
US-09-864-761-41104  
Sequence 41104, Application US/09864761  
Patent No. US20020048763A1  
GENERAL INFORMATION:  
APPLICANT: Penn, Sharron G.  
APPLICANT: Rank, David R.  
APPLICANT: Hanzel, David K.  
APPLICANT: Chen, Wenheng  
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY  
FILE REFERENCE: Aecmica-X-1  
CURRENT APPLICATION NUMBER: US/09/864,761  
CURRENT FILING DATE: 2001-05-23  
PRIOR APPLICATION NUMBER: US 60/180,312  
PRIOR FILING DATE: 2000-02-04  
PRIOR APPLICATION NUMBER: US 60/207,456  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: US 09/632,366  
PRIOR FILING DATE: 2000-08-03  
PRIOR APPLICATION NUMBER: GB 24263.6  
PRIOR FILING DATE: 2000-10-04  
PRIOR APPLICATION NUMBER: US 60/236,359  
PRIOR FILING DATE: 2000-09-27  
PRIOR APPLICATION NUMBER: PCT/US01/006666  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/006667  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/006664  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/006669  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/006665  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/006668  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/006663  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/006662  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/006661  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/006670  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: US 60/234,687  
PRIOR FILING DATE: 2000-09-21  
PRIOR APPLICATION NUMBER: US 09/608,408  
PRIOR FILING DATE: 2000-06-30  
PRIOR APPLICATION NUMBER: US 09/774,203  
PRIOR FILING DATE: 2001-01-29  
NUMBER OF SEQ ID NOS: 49117  
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1

SEQ ID NO 41104

LENGTH: 28  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
OTHER INFORMATION: MAP TO AC022045.2  
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2  
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.2  
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.9  
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.1  
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.6  
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.9  
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.8  
US-09-864-761-41104

Query Match 33.3%; Score 18; DB 9; Length 28;  
Best Local Similarity 28.6%; Pred. No. 4.1e+03;  
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 CXXXXXC 9  
DB 3 CASSSAC 9

## RESULT 20

US-10-697-399-10  
Sequence 10, Application US/10697399  
Publication No. US20040162413A1

GENERAL INFORMATION:  
APPLICANT: Watkins, Jeffrey D.  
APPLICANT: Huse, William D.  
APPLICANT: Vasserot, Alain P.  
APPLICANT: Marquis, David P.  
APPLICANT: Smith, Eric P.  
TITLE OF INVENTION: Methods of Optimizing Antibody Variable Region Binding Affinity  
FILE REFERENCE: AME-08122  
CURRENT APPLICATION NUMBER: US/10/697,399  
CURRENT FILING DATE: 2003-10-30  
NUMBER OF SEQ ID NOS: 50  
SOFTWARE: PatentIn version 3.2  
SEQ ID NO 10  
LENGTH: 33  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Synthetic  
US-10-697-399-10

Query Match 33.3%; Score 18; DB 16; Length 33;  
Best Local Similarity 28.6%; Pred. No. 4.4e+03;  
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 CXXXXXC 9  
DB 15 CATTAAAC 21

## RESULT 21

US-09-816-721-1  
Sequence 1, Application US/09816721  
Patent No. US20020042120A1  
GENERAL INFORMATION:  
APPLICANT: No. US20020042120A1ozymes A/S  
APPLICANT: Andersen, Carsten  
APPLICANT: Nielsen, Bjørne R.  
APPLICANT: Haren, Lubbett D.  
APPLICANT: Dijkstra, Bauke H.  
TITLE OF INVENTION: No. US20020042120A1el Cyclomaltodextrin Glucanotransferase Variant  
FILE REFERENCE: 5347.210-US  
CURRENT APPLICATION NUMBER: US/09/816,721  
CURRENT FILING DATE: 2001-03-22  
NUMBER OF SEQ ID NOS: 2  
SOFTWARE: PatentIn version 3.1

SEQ ID NO 1  
LENGTH: 36  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: synthetic  
NAME/KEY: misc feature  
OTHER INFORMATION: "D196H oligo"  
US-09-816-721-1

Query Match 33.3%; Score 18; DB 9; Length 36;  
Best Local Similarity 28.6%; Pred. No. 4.6e+03;  
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 CXXXXXC 9  
DB 25 CTAAATC 31

## RESULT 22

US-10-029-386-31509  
Sequence 31509, Application US/10029386  
Publication No. US20030194704A1

GENERAL INFORMATION:  
APPLICANT: Penn, Sharron G.  
APPLICANT: Rank, David R.  
APPLICANT: Hanzel, David K.  
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
TITLE OF INVENTION: EXPRESSION ANALYSIS TWO  
FILE REFERENCE: AEOMICA-X-2  
CURRENT APPLICATION NUMBER: US/10/029,386  
CURRENT FILING DATE: 2001-12-20  
NUMBER OF SEQ ID NOS: 34288  
SOFTWARE: Anomax Sequence Listing Engine vers. 1.1  
SEQ ID NO 31509  
LENGTH: 39  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
OTHER INFORMATION: MAP TO AC008456.4  
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.88  
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.87  
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.95  
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 0.62  
US-10-029-386-31509

Query Match 33.3%; Score 18; DB 14; Length 39;  
Best Local Similarity 28.6%; Pred. No. 4.8e+03;  
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 CXXXXXC 9  
DB 19 CSATSTC 25

## RESULT 23

US-08-900-230-58  
Sequence 58, Application US/08900230  
Publication No. US20010009766A1  
GENERAL INFORMATION:  
APPLICANT: Bard, Jonathan A.  
TITLE OF INVENTION: DNA ENCODING GALANN GALR3 RECEPTORS AND  
TITLE OF INVENTION: USES THEREOF  
NUMBER OF SEQUENCES: 59  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Cooper & Dunham LLP  
STREET: 1185 Avenue of The Americas  
CITY: New York  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 11036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/900,230  
FILING DATE: 23-JUL-1997  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: White, John P.  
REGISTRATION NUMBER: 28,678  
REFERENCE/DOCKET NUMBER: 52241-C/JPW/ADM  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-278-0400  
TELEFAX: 212-391-0525  
INFORMATION FOR SEQ ID NO: 58:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 50 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE:  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
US-08-900-230-58

Query Match 33.3%; Score 18; DB 8; Length 50;  
Best Local Similarity 28.6%; Pred. No. 5.5e+03;  
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 CXXXXXC 9  
DB 9 CTAATAC 15

RESULT 24  
US-10-029-386-28972  
; Sequence 28972, Application US/10029386  
; Publication No. US20030194704A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharron G.  
; APPLICANT: Hanzel, David K.  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
; TITLE OF INVENTION: EXPRESSION ANALYSIS TWO  
; FILE REFERENCE: AEMICA-X-2 US/10/029,386  
; CURRENT APPLICATION NUMBER: US/10/029,386  
; CURRENT FILING DATE: 2001-12-20  
; NUMBER OF SEQ ID NOS: 34288  
; SOFTWARE: Annonmax Sequence Listing Engine vers. 1.1  
; SEQ ID NO 28972  
; LENGTH: 58  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: MAP TO CHR22.1  
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.67  
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.1  
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.95  
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 0.61  
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.86  
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.92  
US-10-029-386-28972

Query Match 33.3%; Score 18; DB 14; Length 58;  
Best Local Similarity 28.6%; Pred. No. 5.9e+03;  
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 CXXXXXC 9  
DB 17 CXXXXXC 23

RESULT 25

US-10-437-963-193586  
; Sequence 193586, Application US/10437963  
; Publication No. US20040123343A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Wu, Wei  
; APPLICANT: Boukharov, Andrey A.  
; APPLICANT: Barbazuk, Brad  
; APPLICANT: Li, Ping  
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53221)B  
; CURRENT APPLICATION NUMBER: US/10/437,963  
; CURRENT FILING DATE: 2003-05-14  
; NUMBER OF SEQ ID NOS: 204966  
; SEQ ID NO 193586  
; LENGTH: 66  
; TYPE: PRT  
; ORGANISM: Oryza sativa  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_89711C.1.pap  
US-10-437-963-193586

Query Match 33.3%; Score 18; DB 16; Length 66;  
Best Local Similarity 28.6%; Pred. No. 6.3e+03;  
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 CXXXXXC 9  
DB 35 CSSASSC 41

RESULT 26  
US-10-424-599-202235  
; Sequence 202235, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 202235  
; LENGTH: 67  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_24643C.1.pap  
US-10-424-599-202235

Query Match 33.3%; Score 18; DB 15; Length 67;  
Best Local Similarity 28.6%; Pred. No. 6.4e+03;  
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 CXXXXXC 9  
DB 12 CTTASSC 18

RESULT 27  
US-10-424-599-261219  
; Sequence 261219, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.

APPLICANT: Kovalic David K  
APPLICANT: Zhou Yihua  
APPLICANT: Cao Yongwei  
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
FILE REFERENCE: 38-21(53223)B  
CURRENT APPLICATION NUMBER: US/10/424,599  
CURRENT FILING DATE: 2003-04-28  
NUMBER OF SEQ ID NOS: 285684  
SEQ ID NO 261219  
LENGTH: 70  
TYPE: PRT  
ORGANISM: Glycine max  
FEATURE:  
OTHER INFORMATION: Clone ID: PAT\_MRT3847\_77903C.1.pap  
US-10-424-599-261219

Query Match 33.3%; Score 18; DB 15; Length 70;  
Best Local Similarity 28.6%; Pred. No. 6.5e+03; Indels 0; Gaps 0;  
Matches 2; Conservative 0; Mismatches 5;

QY 3 CXXXXXC 9  
DB 23 CSSASSC 29

RESULT 28  
US-10-011-931-32  
Sequence 32, Application US/10011931  
Publication No. US20030026806A1  
GENERAL INFORMATION:  
APPLICANT: WITTE, ALISON  
APPLICANT: VARNUM, BRIAN C.  
APPLICANT: QIAN, ZUEMING  
APPLICANT: VEZINA, CHRIS  
TITLE OF INVENTION: ANTIBODIES AND OTHER SELECTIVE IL-1 BINDING AGENTS THAT ALLOW BIN  
TITLE OF INVENTION: IL-1 RECEPTOR BUT NOT ACTIVATION THEREOF  
FILE REFERENCE: A-731  
CURRENT APPLICATION NUMBER: US/10/011,931  
CURRENT FILING DATE: 2002-04-01  
PRIOR APPLICATION NUMBER: US 60/244,118  
PRIOR FILING DATE: 2000-10-27  
NUMBER OF SEQ ID NOS: 78  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 32  
LENGTH: 71  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: KAPPA CHAIN CHIMERA  
NAME/KEY: misc feature  
LOCATION: (1)..(1)  
OTHER INFORMATION: At position 1, P = 5' phosphorylated  
US-10-011-931-32

Query Match 33.3%; Score 18; DB 14; Length 71;  
Best Local Similarity 28.6%; Pred. No. 6.6e+03; Indels 0; Gaps 0;  
Matches 2; Conservative 0; Mismatches 5;

QY 3 CXXXXXC 9  
DB 45 CTATTAC 51

RESULT 29  
US-10-269-806-45  
Sequence 45, Application US/10269806  
Publication No. US20030176352A1  
GENERAL INFORMATION:  
APPLICANT: Min, Hosung  
APPLICANT: Sitney, Karen  
APPLICANT: Hartley, Cynthia

TITLE OF INVENTION: Peptides and Related Compounds Having Thrombopoietic Activity  
FILE REFERENCE: A-750  
CURRENT APPLICATION NUMBER: US/10/269,806  
CURRENT FILING DATE: 2002-10-10  
NUMBER OF SEQ ID NOS: 199  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 45  
LENGTH: 71  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Synthesized Peptide Sequence  
US-10-269-806-45

Query Match 33.3%; Score 18; DB 14; Length 71;  
Best Local Similarity 28.6%; Pred. No. 6.6e+03; Indels 0; Gaps 0;  
Matches 2; Conservative 0; Mismatches 5;

QY 3 CXXXXXC 9  
DB 32 CTAAAC 38

RESULT 30  
US-10-269-806-61  
Sequence 61, Application US/10269806  
Publication No. US20030176352A1  
GENERAL INFORMATION:  
APPLICANT: Min, Hosung  
APPLICANT: Sitney, Karen  
APPLICANT: Hartley, Cynthia  
TITLE OF INVENTION: Peptides and Related Compounds Having Thrombopoietic Activity  
FILE REFERENCE: A-750  
CURRENT APPLICATION NUMBER: US/10/269,806  
CURRENT FILING DATE: 2002-10-10  
NUMBER OF SEQ ID NOS: 199  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 61  
LENGTH: 71  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Synthesized Peptide Sequence  
US-10-269-806-61

Query Match 33.3%; Score 18; DB 14; Length 71;  
Best Local Similarity 28.6%; Pred. No. 6.6e+03; Indels 0; Gaps 0;  
Matches 2; Conservative 0; Mismatches 5;

QY 3 CXXXXXC 9  
DB 44 CTTTTTC 50

Search completed: April 27, 2005, 16:48:17  
Job time : 136 secs

**This Page Blank (uspto)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 27, 2005, 16:27:31 ; Search time 164 Seconds  
(without alignments)  
73.107 Million cell updates/sec

Title: US-09-732-348B-34  
Perfect score: 54  
Sequence: 1 XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX 31

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 150 summaries

Database : A\_Geneseq\_16Dec04.\*

- 1: Geneseqp1980s.\*
- 2: Geneseqp1990s.\*
- 3: Geneseqp2000s.\*
- 4: Geneseqp2001s.\*
- 5: Geneseqp2002s.\*
- 6: Geneseqp2003as.\*
- 7: Geneseqp2003bs.\*
- 8: Geneseqp2004s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	19	35.2	103	7	ABO23576 Mycoplasma
2	18	33.3	10	2	Aaw65945 Molecule
3	18	33.3	10	4	Aag96291 Human com
4	18	33.3	10	4	Aag96207 Human com
5	18	33.3	10	4	Aag96231 Human com
6	18	33.3	10	4	Aag96205 Human com
7	18	33.3	10	4	Aag96349 Human com
8	18	33.3	10	4	Aag96191 Human com
9	18	33.3	10	4	Aag96339 Human com
10	18	33.3	10	4	Aag96363 Human com
11	18	33.3	10	4	Aag96333 Human com
12	18	33.3	10	4	Aag96193 Human com
13	18	33.3	12	2	Aaw54432 Human P51
14	18	33.3	12	3	ABO8422 Antigen p
15	18	33.3	12	8	ADQ96644 Human com
16	18	33.3	15	2	Aar97928 Japan ced
17	18	33.3	15	2	Aar97929 Japan ced
18	18	33.3	18	6	ABP83105 G protein
19	18	33.3	20	2	Aaw42192 T-cell ep
20	18	33.3	25	7	ADM10961 Human ova
21	18	33.3	25	8	ADJ11291 Rabbit an
22	18	33.3	28	4	ABb42303 Peptide #
23	18	33.3	28	4	ABb42303 Peptide #
24	18	33.3	28	4	ABb25806 Protein #
25	18	33.3	28	4	Aam76003 Human bon

26	18	33.3	28	4	AAM63189 Human bra
27	18	33.3	28	4	ABG57729 Human liv
28	18	33.3	28	5	ABG45410 Human pep
29	18	33.3	37	8	ABO57875 Human gen
30	18	33.3	57	4	ABG05767 Novel hum
31	18	33.3	58	8	ABO55338 Human gen
32	18	33.3	60	2	AAR98207 Nucleotid
33	18	33.3	62	8	ADP31006 Human sec
34	18	33.3	69	4	AAU62802 Propionib
35	18	33.3	69	6	ARM59321 Propionib
36	18	33.3	79	4	AAO11175 Human pol
37	18	33.3	81	2	RAY48268 Human pro
38	18	33.3	86	3	AAB56561 Human pro
39	18	33.3	88	3	AAAG21579 Arabidops
40	18	33.3	88	3	AAAG52968 Arabidops
41	18	33.3	90	2	RAY59864 Human nor
42	18	33.3	94	3	AAAB41818 Human ORF
43	18	33.3	99	8	ADP31678 Human sec
44	18	33.3	104	3	AAAB40237 Human ORF
45	18	33.3	104	5	ABP31494 Human ORF
46	18	33.3	113	8	ADP31617 Human sec
47	18	33.3	114	7	ADM10958 Human ova
48	18	33.3	114	8	ADJ11288 Human ova
49	18	33.3	115	7	ADM10959 Human ova
50	18	33.3	115	8	ADJ11289 Human ova
51	18	33.3	120	5	ABP10204 Human ORF
52	18	33.3	121	2	AAAS5794 Herpesvir
53	18	33.3	121	2	AAAR86314 Herpesvir
54	18	33.3	125	7	ABO78980 Pseudomon
55	18	33.3	126	4	ABG27631 Novel hum
56	18	33.3	127	8	ADP31622 Human sec
57	18	33.3	130	8	ADR42993 IPT-like
58	18	33.3	131	7	ADM05721 Human pro
59	18	33.3	132	7	ABO73810 Pseudomon
60	18	33.3	133	4	AAU51326 Propionib
61	18	33.3	133	4	AAU64082 Propionib
62	18	33.3	133	4	ABG01178 Novel hum
63	18	33.3	133	6	ABM47845 Propionib
64	18	33.3	133	6	ABM60601 Propionib
65	18	33.3	134	5	ABU51077 Helicobac
66	18	33.3	134	7	ABO76323 Pseudomon
67	18	33.3	135	3	AAAB01404 Rat Ly-68
68	18	33.3	137	8	ADM10953 Human O59
69	18	33.3	137	8	ADJ11283 Human ova
70	18	33.3	137	8	ADM43544 Human ova
71	18	33.3	138	3	AAAB42027 Human ORF
72	18	33.3	138	7	ABO67837 Pseudomon
73	18	33.3	139	7	ABO80767 Pseudomon
74	18	33.3	141	5	AAU83599 Human PRO
75	18	33.3	141	5	ABB09417 Ovarian c
76	18	33.3	141	5	AAO18096 Human ova
77	18	33.3	141	6	ABR58665 Human can
78	18	33.3	141	6	ABU80746 Human PRO
79	18	33.3	141	6	ABO33712 Novel hum
80	18	33.3	141	6	ABP97182 Tumour-as
81	18	33.3	141	6	ABU82055 Novel hum
82	18	33.3	141	6	ABJ72235 Human PRO
83	18	33.3	141	6	ABJ72363 Human PRO
84	18	33.3	141	6	ABO34258 Human sec
85	18	33.3	141	7	ABJ72065 Human mem
86	18	33.3	141	7	ABD83506 Novel hum
87	18	33.3	141	7	ABD80612 Novel hum
88	18	33.3	141	7	ADB73153 Novel hum
89	18	33.3	141	7	ADB78235 Novel hum
90	18	33.3	141	7	ADB84883 Human PRO
91	18	33.3	141	7	ADB77989 Novel hum
92	18	33.3	141	7	ADB87055 Human PRO
93	18	33.3	141	7	ADB84637 Human PRO
94	18	33.3	141	7	ADB83752 Novel hum
95	18	33.3	141	7	ADB72907 Novel hum
96	18	33.3	141	7	ADC36745 Human PRO
97	18	33.3	141	7	ADC21735 Human PRO
98	18	33.3	141	7	ADC49766 Novel hum

99 18 33.3 141 7 ADC48965  
 100 18 33.3 141 7 ADC49482  
 101 18 33.3 141 7 ADC47343  
 102 18 33.3 141 7 ADC47088  
 103 18 33.3 141 7 ADC77963  
 104 18 33.3 141 7 ADD06198  
 105 18 33.3 141 7 ADC77717  
 106 18 33.3 141 7 ADC50680  
 107 18 33.3 141 7 ADD50926  
 108 18 33.3 141 7 ADD50407  
 109 18 33.3 141 7 ADD50161  
 110 18 33.3 141 7 ADD51172  
 111 18 33.3 141 7 ADI21041  
 112 18 33.3 141 7 ADM03955  
 113 18 33.3 141 7 ADM10884  
 114 18 33.3 141 8 ADC48719  
 115 18 33.3 141 8 ADE20890  
 116 18 33.3 141 8 ADE05734  
 117 18 33.3 141 8 ADD74963  
 118 18 33.3 141 8 ADD75709  
 119 18 33.3 141 8 ADD84941  
 120 18 33.3 141 8 ADD86767  
 121 18 33.3 141 8 ADE20644  
 122 18 33.3 141 8 ADE38941  
 123 18 33.3 141 8 ADE05488  
 124 18 33.3 141 8 ADD73473  
 125 18 33.3 141 8 ADD78313  
 126 18 33.3 141 8 ADE21136  
 127 18 33.3 141 8 ADD77251  
 128 18 33.3 141 8 ADE20398  
 129 18 33.3 141 8 ADD75463  
 130 18 33.3 141 8 ADD73979  
 131 18 33.3 141 8 ADD74225  
 132 18 33.3 141 8 ADD75955  
 133 18 33.3 141 8 ADD85447  
 134 18 33.3 141 8 ADE04996  
 135 18 33.3 141 8 ADD75209  
 136 18 33.3 141 8 ADD76753  
 137 18 33.3 141 8 ADD86521  
 138 18 33.3 141 8 ADD77989  
 139 18 33.3 141 8 ADD77497  
 140 18 33.3 141 8 ADD77743  
 141 18 33.3 141 8 ADD85201  
 142 18 33.3 141 8 ADD73733  
 143 18 33.3 141 8 ADD74471  
 144 18 33.3 141 8 ADD76999  
 145 18 33.3 141 8 ADD85693  
 146 18 33.3 141 8 ADE05242  
 147 18 33.3 141 8 ADD74717  
 148 18 33.3 141 8 ADG05529  
 149 18 33.3 141 8 ADG27083  
 150 18 33.3 141 8 ADG11146

## ALIGNMENTS

RESULT 1  
 ABO23576  
 ID ABO23576 standard; protein; 103 AA.  
 XX  
 AC ABO23576;  
 XX  
 DT 04-SEP-2003 (first entry)  
 XX  
 DE Mycoplasma pneumoniae outlier protein #4.  
 KW Candidate protein identification; pathogen; anti-infective;  
 KW outlier protein; virulence protein; antigen; drug target protein;  
 KW pathogenic organism; antimicrobial.  
 XX Mycoplasma pneumoniae.  
 XX

PN US2003039963-A1.  
 XX 27-FEB-2003.  
 PD 30-MAR-2001; 2001US-00820843.  
 XX 30-MAR-2001; 2001US-00820843.  
 XX (BRAH/) BRAHMACHARI S K.  
 PA (RAWA/) RAMACHANDRAN S.  
 PA (NAND/) NANDI T.  
 PA (BHIM/) BHIMARAO C.  
 XX  
 PI Brahmachari SK, Ramachandran S, Nandi T, Bhimmarao C;  
 XX WPI; 2003-492159/46.  
 DR  
 XX  
 PT Identifying candidate proteins useful as anti-infectives involves  
 PT matching outlier protein sequences with protein sequences in databases.  
 XX  
 PS Example 7; Page 78; 117pp; English.  
 XX  
 CC The present invention relates to a method for identifying candidate  
 CC proteins in pathogens useful as anti-infectives. The invention discloses  
 CC a computational method which involves the calculation of several sequence  
 CC attributes and their subsequent analysis results in the identification  
 CC of outlier proteins in different pathogens. The method is useful for the  
 CC identification of outlier proteins (e.g. virulence proteins, antigens or  
 CC proteins used as drug targets) in pathogenic organisms. The method of the  
 CC invention provides reproducible results as it does not depend on the  
 CC variable biochemical characterisation of proteins. ABO23500-ABO23617  
 CC represent outlier proteins identified from different pathogenic organisms  
 XX  
 SQ Sequence 103 AA;  
 Query Match 35.2%; Score 19; DB 7; Length 103;  
 Best Local Similarity 13.6%; Pred. No. 1.2e+04;  
 Matches 3; Conservative 0; Mismatches 19; Indels 0; Gaps 0;  
 QY 3 CXXXXXXCXXXXXXCXXXXXXH 24  
 DB 53 CSSISFCSLASSARLYRSSH 74  
 RESULT 2  
 AAW65945  
 ID AAW65945 standard; peptide; 10 AA.  
 XX  
 AC AAW65945;  
 XX  
 DT 12-NOV-1998 (first entry)  
 XX  
 DE Molecule V.  
 XX  
 KW Neurotrophin; cyclic; bicyclic; disulphide; nerve growth factor; NGF;  
 KW BDNF; NT-3; conformation; promoter.  
 XX  
 OS Synthetic.  
 OS Mammalia.  
 XX  
 PN CA2205045-AA.  
 XX  
 PD 12-MAY-1998.  
 XX  
 PF 09-MAY-1997; 97CA-02205045.  
 XX  
 PR 12-NOV-1996; 96CA-02190296.  
 XX  
 PA (TOOH ) UNIV QUEENS KINGSTON.  
 XX  
 PI Riopelle RJ, Dory MI, Shamovsky IL, Weaver DF, Ross GM;  
 XX WPI; 1998-457650/40.  
 DR





RESULT 8  
 AAG96191  
 ID AAG96191 standard; peptide; 10 AA.  
 XX  
 AC AAG96191;  
 XX  
 DT 18-SEP-2001 (first entry)  
 XX  
 DE Human complementary peptide, SEQ ID NO: 2385.  
 XX  
 KW Human; complementary peptide; ligand; drug discovery; drug design.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200142277-A2.  
 XX  
 PD 14-JUN-2001.  
 XX  
 PF 13-DEC-2000; 2000WO-GB004776.  
 XX  
 PR 13-DEC-1999; 99GB-00029464.  
 XX  
 PA (PROT-) PROTEOM LTD.  
 XX  
 PI Roberts GW, Heal JR;  
 XX  
 DR WPI; 2001-408419/43.  
 XX  
 PT A set of peptide ligands consisting of specific complementary peptides to  
 PT proteins encoded by genes of the human genome, useful in an assay for  
 PT screening and identifying of one or more novel peptides which are drug  
 PT candidates or pro-drugs.  
 XX  
 PS Example 4; Page 383; 646pp; English.  
 XX  
 CC The invention relates to a set of complementary peptide ligands generated  
 CC from the human genome. The complementary peptides interact with their  
 CC relevant target proteins encoded in the human genome. They can be used as  
 CC reagents in drug discovery and as lead ligands to facilitate drug design  
 CC and development. The present sequence is a complementary peptide provided  
 CC in the specification  
 CC  
 SQ Sequence 10 AA;  
 XX  
 Query Match 33.3%; Score 18; DB 4; Length 10;  
 Best Local Similarity 28.6%; Pred. NO. 6.1e+03;  
 Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
 OY 3 CXXXXXC 9  
 DB 3 CSSSTTC 9  
 XX  
 RESULT 9  
 AAG96339  
 ID AAG96339 standard; peptide; 10 AA.  
 XX  
 AC AAG96339;  
 XX  
 DT 18-SEP-2001 (first entry)  
 XX  
 DE Human complementary peptide, SEQ ID NO: 2533.  
 XX  
 KW Human; complementary peptide; ligand; drug discovery; drug design.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200142277-A2.  
 XX  
 PD 14-JUN-2001.  
 XX  
 PT A set of peptide ligands consisting of specific complementary peptides to  
 PT proteins encoded by genes of the human genome, useful in an assay for  
 PT screening and identifying of one or more novel peptides which are drug  
 PT candidates or pro-drugs.  
 XX  
 PS Example 4; Page 383; 646pp; English.  
 XX  
 CC The invention relates to a set of complementary peptide ligands generated  
 CC from the human genome. The complementary peptides interact with their  
 CC relevant target proteins encoded in the human genome. They can be used as  
 CC reagents in drug discovery and as lead ligands to facilitate drug design  
 CC and development. The present sequence is a complementary peptide provided  
 CC in the specification  
 CC  
 SQ Sequence 10 AA;  
 XX  
 Query Match 33.3%; Score 18; DB 4; Length 10;  
 Best Local Similarity 28.6%; Pred. NO. 6.1e+03;  
 Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
 OY 3 CXXXXXC 9  
 DB 3 CSSSTTC 9  
 XX  
 RESULT 9  
 AAG96339  
 ID AAG96339 standard; peptide; 10 AA.  
 XX  
 AC AAG96339;  
 XX  
 DT 18-SEP-2001 (first entry)  
 XX  
 DE Human complementary peptide, SEQ ID NO: 2533.  
 XX  
 KW Human; complementary peptide; ligand; drug discovery; drug design.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200142277-A2.  
 XX  
 PD 14-JUN-2001.  
 XX

PF 13-DEC-2000; 2000WO-GB004776.  
 XX  
 PR 13-DEC-1999; 99GB-00029464.  
 XX  
 PA (PROT-) PROTEOM LTD.  
 XX  
 PI Roberts GW, Heal JR;  
 XX  
 DR WPI; 2001-408419/43.  
 XX  
 PT A set of peptide ligands consisting of specific complementary peptides to  
 PT proteins encoded by genes of the human genome, useful in an assay for  
 PT screening and identifying of one or more novel peptides which are drug  
 PT candidates or pro-drugs.  
 XX  
 PS Example 4; Page 404; 646pp; English.  
 XX  
 CC The invention relates to a set of complementary peptide ligands generated  
 CC from the human genome. The complementary peptides interact with their  
 CC relevant target proteins encoded in the human genome. They can be used as  
 CC reagents in drug discovery and as lead ligands to facilitate drug design  
 CC and development. The present sequence is a complementary peptide provided  
 CC in the specification  
 CC  
 SQ Sequence 10 AA;  
 XX  
 Query Match 33.3%; Score 18; DB 4; Length 10;  
 Best Local Similarity 28.6%; Pred. NO. 6.1e+03;  
 Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
 OY 3 CXXXXXC 9  
 DB 2 CSSSTTC 8  
 XX  
 RESULT 10  
 AAG96363  
 ID AAG96363 standard; peptide; 10 AA.  
 XX  
 AC AAG96363;  
 XX  
 DT 18-SEP-2001 (first entry)  
 XX  
 DE Human complementary peptide, SEQ ID NO: 2557.  
 XX  
 KW Human; complementary peptide; ligand; drug discovery; drug design.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200142277-A2.  
 XX  
 PD 14-JUN-2001.  
 XX  
 PF 13-DEC-2000; 2000WO-GB004776.  
 XX  
 PR 13-DEC-1999; 99GB-00029464.  
 XX  
 PA (PROT-) PROTEOM LTD.  
 XX  
 PI Roberts GW, Heal JR;  
 XX  
 DR WPI; 2001-408419/43.  
 XX  
 PT A set of peptide ligands consisting of specific complementary peptides to  
 PT proteins encoded by genes of the human genome, useful in an assay for  
 PT screening and identifying of one or more novel peptides which are drug  
 PT candidates or pro-drugs.  
 XX  
 PS Example 4; Page 408; 646pp; English.  
 XX  
 CC The invention relates to a set of complementary peptide ligands generated  
 CC from the human genome. The complementary peptides interact with their  
 CC relevant target proteins encoded in the human genome. They can be used as

CC reagents in drug discovery and as lead ligands to facilitate drug design  
CC and development. The present sequence is a complementary peptide provided  
CC in the specification

XX SQ Sequence 10 AA;

Query Match 33.3%; Score 18; DB 4; Length 10;  
Best Local Similarity 28.6%; Pred. No. 6.1e+03;  
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 3 CXXXXXC 9  
Db 2 CSSSTTC 8

RESULT 11  
AAG96333  
ID AAG96333 standard; peptide; 10 AA.

XX AC AAG96333;

XX DT 18-SEP-2001 (first entry)

XX DE Human complementary peptide, SEQ ID NO: 2527.

XX KW Human; complementary peptide; ligand; drug discovery; drug design.

XX OS Homo sapiens.

XX PN WO200142277-A2.

XX PD 14-JUN-2001.

XX PF 13-DEC-2000; 2000WO-GB004776.

XX PR 13-DEC-1999; 99GB-00029464.

XX PA (PROT-) PROTEOM LTD.

XX PI Roberts GW, Heal JR;

XX DR WPI; 2001-408419/43.

XX PT A set of peptide ligands consisting of specific complementary peptides to  
PT proteins encoded by genes of the human genome, useful in an assay for  
PT screening and identifying of one or more novel peptides which are drug  
PT candidates or pro-drugs.

XX PS Example 4; Page 403; 646pp; English.

XX CC The invention relates to a set of complementary peptide ligands generated  
CC from the human genome. The complementary peptides interact with their  
CC relevant target proteins encoded in the human genome. They can be used as  
CC reagents in drug discovery and as lead ligands to facilitate drug design  
CC and development. The present sequence is a complementary peptide provided  
CC in the specification

XX SQ Sequence 10 AA;

Query Match 33.3%; Score 18; DB 4; Length 10;  
Best Local Similarity 28.6%; Pred. No. 6.1e+03;  
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 3 CXXXXXC 9  
Db 2 CSSSTTC 8

RESULT 12

AAG96193

ID AAG96193 standard; peptide; 10 AA.

XX AC AAG96193;

XX DT 18-SEP-2001 (first entry)  
XX DE Human complementary peptide, SEQ ID NO: 2387.

XX KW Human; complementary peptide; ligand; drug discovery; drug design.

XX OS Homo sapiens.

XX PN WO200142277-A2.

XX PD 14-JUN-2001.

XX PF 13-DEC-2000; 2000WO-GB004776.

XX PR 13-DEC-1999; 99GB-00029464.

XX PA (PROT-) PROTEOM LTD.

XX PI Roberts GW, Heal JR;

XX DR WPI; 2001-408419/43.

XX PT A set of peptide ligands consisting of specific complementary peptides to  
PT proteins encoded by genes of the human genome, useful in an assay for  
PT screening and identifying of one or more novel peptides which are drug  
PT candidates or pro-drugs.

XX PS Example 4; Page 384; 646pp; English.

XX CC The invention relates to a set of complementary peptide ligands generated  
CC from the human genome. The complementary peptides interact with their  
CC relevant target proteins encoded in the human genome. They can be used as  
CC reagents in drug discovery and as lead ligands to facilitate drug design  
CC and development. The present sequence is a complementary peptide provided  
CC in the specification

XX SQ Sequence 10 AA;

Query Match 33.3%; Score 18; DB 4; Length 10;  
Best Local Similarity 28.6%; Pred. No. 6.1e+03;  
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 3 CXXXXXC 9  
Db 4 CSSSTTC 10

RESULT 13

AAW54432

ID AAW54432 standard; peptide; 12 AA.

XX AC AAW54432;

XX DT 15-SEP-1998 (first entry)

XX DE Human PS112 protein synthetic peptide #7.

XX KW Prostate; disease; PS112 gene; detection; diagnosis; cancer; treatment;  
KW antibody.

XX OS Homo sapiens.

XX OS Synthetic.

XX PN WO9815657-A1.

XX PD 16-APR-1998.

XX PF 08-OCT-1997; 97WO-US018290.

XX PR 08-OCT-1996; 96US-00727688.

XX PA (ABRO ) ABBOTT LAB.

XX Cohen M, Friedman PN, Gordon J, Hodges SC, Klass MR;  
 PI Kratochvil JD, Roberts-Rapp L, Russell JC, Stroupe SD;  
 DR WPI; 1998-240838/21.  
 XX Detecting a target PS112 polynucleotide - used for diagnosing prostate  
 PT cancer.  
 PT  
 XX  
 PS Example 10; Page 92; 104pp; English.  
 XX  
 CC AAW54426-W54433 represent synthetic peptide fragments of designed from  
 CC the human PS112 protein isolated from a prostate library. These peptides  
 CC are used in a novel method of detecting the presence of a target PS112  
 CC polynucleotide in a test sample. The method can also be used to detect  
 CC mRNA of PS112 in a test sample. The method can be used for diagnosis of  
 CC prostate cancer, as the presence of PS112 is an indicator of prostate  
 CC cancer. Antibodies against the polypeptides may be used as markers, or to  
 CC treat prostate cancer  
 XX  
 SQ Sequence 12 AA;  
 Query Match 33.3%; Score 18; DB 2; Length 12;  
 Best Local Similarity 28.6%; Pred. No. 6.7e+03;  
 Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
 QY 3 CXXXXXC 9  
 DB 1 CSSSSSC 7  
 RESULT 14  
 ID AAB08422 standard; peptide; 12 AA.  
 XX  
 AC AAB08422;  
 XX  
 DT 20-DEC-2000 (first entry)  
 XX  
 DE Antigen peptide derived from prostate cancer associated protein PS112.  
 XX  
 KW Prostate cancer associated gene; PS112; prostate disease;  
 KW prostate cancer; tumour; metastasis.  
 XX  
 OS Synthetic.  
 OS Homo sapiens.  
 XX  
 PN US6110675-A.  
 XX  
 PD 29-AUG-2000.  
 XX  
 PF 08-OCT-1997; 97US-00946869.  
 XX  
 PR 08-OCT-1996; 96US-00727688.  
 XX  
 PA (ABBO ) ABBOTT LAB.  
 XX  
 PI Friedman PN, Gordon J, Hodges SC, Klass MR, Cohen M;  
 PI Roberts-Rapp L, Russell JC, Stroupe SD, Yu H, Kratochvil JD;  
 XX  
 DR WPI; 2000-571422/53.  
 XX  
 PT Novel methods for diagnosing prostate cancer by contacting test sample  
 PT with target specific polynucleotide and detecting prostate cancer  
 PT associated polynucleotides.  
 XX  
 PS Example 10; Col 77-78; 50pp; English.  
 XX  
 CC AAB08416-23 represent antigenic peptides derived from a protein which is  
 CC encoded by a human prostate cancer associated gene, designated PS112. The  
 CC peptides are used to raise antibodies. PS112 sequences are useful for  
 CC detecting, diagnosing, staging, monitoring, prognosticating, preventing,  
 CC treating, or determining the predisposition of an individual to disease

CC and conditions of the prostate, such as prostate cancer, tumours and  
 CC metastases  
 XX  
 SQ Sequence 12 AA;  
 Query Match 33.3%; Score 18; DB 3; Length 12;  
 Best Local Similarity 28.6%; Pred. No. 6.7e+03;  
 Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
 QY 3 CXXXXXC 9  
 DB 1 CSSSSSC 7  
 RESULT 15  
 ID ADQ96644 standard; peptide; 12 AA.  
 XX  
 AC ADQ96644;  
 XX  
 DT 23-SEP-2004 (first entry)  
 XX  
 DE Human PS112 epitope seqid 32.  
 XX  
 KW cytostatic; gene therapy; PS112; recombinant expression system;  
 KW PS112 epitope; prostate disease; tumours; metastasis; predisposition;  
 KW prostate cancer; epitope.  
 XX  
 OS Synthetic.  
 XX  
 PN US2004121397-A1.  
 XX  
 PD 24-JUN-2004.  
 XX  
 PF 22-JAN-2004; 2004US-00763992.  
 XX  
 PR 08-OCT-1996; 96US-00727688.  
 PR 08-OCT-1997; 97US-00946869.  
 PR 15-OCT-1999; 99US-00418887.  
 XX  
 PA (COHE/) COHEN M.  
 PA (FRIE/) FRIEDMAN P N.  
 PA (GORD/) GORDON J.  
 PA (HODG/) HODGES S C.  
 PA (KLAS/) KLASS M R.  
 PA (KRAT/) KRATOCHVIL J D.  
 PA (ROBE/) ROBERTS-RAPP L.  
 PA (RUSS/) RUSSELL J C.  
 PA (STRO/) STROUPE S D.  
 PA (YUHH/) YU H.  
 XX  
 PI Cohen M, Friedman PN, Gordon J, Hodges SC, Klass MR;  
 PI Kratochvil JD, Roberts-Rapp L, Russell JC, Stroupe SD, Yu H;  
 XX  
 DR WPI; 2004-479676/45.  
 XX  
 PT Detecting a target PS112 polynucleotide, useful in diagnosing, staging,  
 PT monitoring, prognosticating, preventing and treating prostate cancer,  
 PT comprises contacting the test sample with PS112-specific polynucleotide.  
 XX  
 PS Example 10; SEQ ID NO 32; 53pp; English.  
 XX  
 CC The invention describes a method of detecting the presence of a target  
 CC PS112 polynucleotide in a test sample. The method comprises: contacting  
 CC the test sample with at least one PS112-specific polynucleotide or its  
 CC complement; and detecting the presence of the target PS112 polynucleotide  
 CC in the test sample, where the PS112-specific polynucleotide has at least  
 CC 50% identity to a polynucleotide comprising a sequence of 367, 214, 205,  
 CC 256, 246, 277, 251, 223, 2393, or 1297 bp (SEQ ID NOS: 1-10) or their  
 CC fragments or complements. Also described are: detecting mRNA of PS112 in  
 CC a test sample; a test kit, useful for detecting PS112 polynucleotide in  
 CC test sample; a purified polynucleotide or fragment derived from a PS112  
 CC gene; a recombinant expression system comprising a nucleic acid sequence

CC that includes an open reading frame derived from P8112 operably linked to  
 CC a control sequence compatible with a desired host, where the nucleic acid  
 CC sequence has at least 50% identity to a sequence of SEQ ID NOS: 1-10, or  
 CC their fragments or complements; a cell transfected with the recombinant  
 CC expression system or with a nucleic acid sequence encoding at least one  
 CC P8112 epitope, where the nucleic acid sequence comprises SEQ ID NOS: 1-  
 CC 10, or their fragments or complements; a composition of matter comprising  
 CC a P8112 polynucleotide or its fragment, where the polynucleotide has at  
 CC least 50% identity to a sequence of SEQ ID NOS: 2-10, or their  
 CC complements, or has at least 50% identity with fragments of a  
 CC polynucleotide of SEQ ID NOS: 4-8; and a gene or its fragment comprising  
 CC DNA having at least 50% identity with SEQ ID NOS: 9 or 10. The method is  
 CC useful for detecting the presence of a target P8112 polynucleotide in a  
 CC test sample. The methods, test kit, polynucleotides and polypeptides, and  
 CC antibodies are useful in detecting, diagnosing, staging, monitoring, or  
 CC prognosticating, preventing and treating prostate diseases, tumours or  
 CC metastases or in determining the predisposition of an individual to  
 CC diseases and conditions of the prostate, e.g. prostate cancer. This is  
 CC the amino acid sequence of a P8112 epitope used in the creation of anti-  
 CC P8112-antibodies.

XX SQ Sequence 12 AA;

Query Match 33.3%; Score 18; DB 8; Length 12;  
 Best Local Similarity 28.6%; Pred. No. 6.7e+03;  
 Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 3 CXXXXXC 9  
 Db 1 CXXXXXC 7

RESULT 16

AAR97928  
 ID AAR97928 standard; peptide; 15 AA.

XX AC AAR97928;

XX DT 16-AUG-1996 (first entry)

XX DE Japan cedar pollen mature allergen Cry j II amino acids 286-300.

XX KW Allergen; epitope; overlapping peptide; Cry j II; cedar pollen;

XX KW Sugi pollinosis; diagnosis; treatment.

XX OS Cryptomeria japonica.

XX PN JP08047392-A.

XX PD 20-FEB-1996.

XX PF 07-NOV-1994; 94JP-00297840.

XX PR 05-NOV-1993; 93JP-00276773.

XX PR 26-MAY-1994; 94JP-00134868.

XX PA (MEIP ) MEIJI MILK PROD CO LTD.

XX DR WPI; 1996-166249/17.

XX PT Japan cedar pollen allergen Cry j II epitope - comprises at least part of  
 PT specified 460 aminoacid protein.

XX PS Disclosure; Fig 5; 17pp; Japanese.

XX CC AAR97871-R97960 are overlapping peptides used for the epitope mapping of  
 CC the Japan cedar pollen allergen Cry j II. Cry j II and allergenic  
 CC peptides of it are useful in the diagnosis, prevention and treatment of  
 CC Sugi pollinosis, the allergic reaction to Japan cedar pollen. Significant  
 CC regions of the allergen were identified using the overlapping peptides of  
 CC the full epitope derived from a Cry j II antigen-specific T cell line.  
 CC Amino acids 66-80 (AAR97884) and 186-200 (R978908) of the full mature 460  
 CC amino acid allergen are the most allergenic of the 90 peptides tested

XX SQ Sequence 15 AA;

Query Match 33.3%; Score 18; DB 2; Length 15;  
 Best Local Similarity 28.6%; Pred. No. 7.4e+03;  
 Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 3 CXXXXXC 9  
 Db 8 CTSASAC 14

RESULT 17

AAR97929  
 ID AAR97929 standard; peptide; 15 AA.

XX AC AAR97929;

XX DT 16-AUG-1996 (first entry)

XX DE Japan cedar pollen mature allergen Cry j II amino acids 291-310.

XX KW Allergen; epitope; overlapping peptide; Cry j II; cedar pollen;

XX KW Sugi pollinosis; diagnosis; treatment.

XX OS Cryptomeria japonica.

XX PN JP08047392-A.

XX PD 20-FEB-1996.

XX PF 07-NOV-1994; 94JP-00297840.

XX PR 05-NOV-1993; 93JP-00276773.

XX PR 26-MAY-1994; 94JP-00134868.

XX PA (MEIP ) MEIJI MILK PROD CO LTD.

XX DR WPI; 1996-166249/17.

XX PT Japan cedar pollen allergen Cry j II epitope - comprises at least part of  
 PT specified 460 aminoacid protein.

XX PS Disclosure; Fig 5; 17pp; Japanese.

XX CC AAR97871-R97960 are overlapping peptides used for the epitope mapping of  
 CC the Japan cedar pollen allergen Cry j II. Cry j II and allergenic  
 CC peptides of it are useful in the diagnosis, prevention and treatment of  
 CC Sugi pollinosis, the allergic reaction to Japan cedar pollen. Significant  
 CC regions of the allergen were identified using the overlapping peptides of  
 CC the full epitope derived from a Cry j II antigen-specific T cell line.  
 CC Amino acids 66-80 (AAR97884) and 186-200 (R978908) of the full mature 460  
 CC amino acid allergen are the most allergenic of the 90 peptides tested

XX SQ Sequence 15 AA;

Query Match 33.3%; Score 18; DB 2; Length 15;  
 Best Local Similarity 28.6%; Pred. No. 7.4e+03;  
 Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 3 CXXXXXC 9  
 Db 3 CTSASAC 9

RESULT 18

ABP83105  
 ID ABP83105 standard; peptide; 18 AA.

XX AC ABP83105;

XX DT 04-MAR-2003 (first entry)

XX XX



DE G protein-coupled receptor (GPCR) antigenic peptide SEQ ID NO:1778.  
 XX  
 KW G protein-coupled receptor; GPCR; antigenic peptide; gene therapy;  
 KW G protein-coupled receptor modulator; antibody; immune-related disease;  
 KW growth-related disease; cell regeneration-related disease; AIDS; cancer;  
 KW immunological-related cell proliferative disease; autoimmune disease;  
 KW Alzheimer's disease; atherosclerosis; infection; osteoarthritis; allergy;  
 KW osteoporosis; cardiomyopathy; inflammation; Crohn's disease; diabetes;  
 KW graft versus host disease; Parkinson's disease; multiple sclerosis; pain;  
 KW psoriasis; anxiety; depression; schizophrenia; dementia; memory loss;  
 KW mental retardation; epilepsy; asthma; tuberculosis; obesity; nausea;  
 KW hypertension; hypotension; renal disorder; rheumatoid arthritis; trauma;  
 KW ulcer.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200261087-A2.  
 XX  
 PD 08-AUG-2002.  
 XX  
 PF 19-DEC-2001; 2001WO-US050107.  
 XX  
 PR 19-DEC-2000; 2000US-0257144P.  
 XX  
 PA (LIFE-) LIFESPAN BIOSCIENCES INC.  
 XX  
 PI Burmer GC, Roush CL, Brown JP;  
 XX  
 PI WPI; 2003-046718/04.  
 XX  
 DR  
 XX New isolated antigenic peptides e.g., for G protein-coupled receptors  
 PT (GPCR), useful for diagnosing and designing drugs for treating conditions  
 PT in which GPCRs are involved, e.g. AIDS, Alzheimer's disease, cancer or  
 PT autoimmune diseases.  
 XX  
 PS Claim 1; Fig 2; 523pp; English.  
 XX  
 CC The present invention describes antigenic peptides (I) comprising: (a)  
 CC any one of 1601 sequences (see ABP82019 to ABP83619) of 12-24 amino  
 CC acids. Also described: (1) an assay for the detection of a particular G  
 CC protein-coupled receptor (GPCR) or a candidate polypeptide in a sample;  
 CC and (2) an isolated antibody having high specificity and high affinity or  
 CC avidity for a particular GPCR. (I) can be used as GPCR modulators and in  
 CC gene therapy. The antigenic peptides for GPCRs are useful in detecting an  
 CC antibody against a particular GPCR, and in the production of specific  
 CC antibodies. The peptides and antibodies are also useful for detecting the  
 CC presence or absence of corresponding GPCRs. The antigenic peptides for  
 CC GPCRs and antibodies are useful for diagnosing and designing drugs for  
 CC treating immune-related diseases, growth-related diseases, cell  
 CC regeneration-related disease, immunological-related cell proliferative  
 CC diseases, or autoimmune diseases, e.g. AIDS, Alzheimer's disease,  
 CC atherosclerosis, bacterial, fungal, protozoan or viral infections,  
 CC osteoarthritis, osteoporosis, cancer, cardiomyopathy, chronic and acute  
 CC inflammation, allergies, Crohn's disease, diabetes, graft versus host  
 CC disease, Parkinson's disease, multiple sclerosis, pain, psoriasis,  
 CC anxiety, depression, schizophrenia, dementia, mental retardation, memory  
 CC loss, epilepsy, asthma, tuberculosis, obesity, nausea, hypertension,  
 CC hypotension, renal disorders, rheumatoid arthritis, trauma, ulcers, or  
 CC any other disorder in which GPCRs are involved. The antibodies may be  
 CC used in immunoassays and immunodiagnosis. ABZ42523 to ABZ42869 encode  
 CC GPCR proteins given in ABP81675 to ABP82018, which are used in the  
 CC exemplification of the present invention  
 XX  
 SQ Sequence 18 AA;

Query Match 33.3%; Score 18; DB 6; Length 18;  
 Best Local Similarity 28.6%; Pred. No. 8.1e+03;  
 Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
 QY 3 CXXXXXC 9  
 DB 9 CASSSC 15

RESULT 19  
 AAW42192  
 ID AAW42192 standard; peptide; 20 AA.  
 XX  
 AC AAW42192;  
 XX  
 DT 27-AUG-2003 (revised)  
 DT 25-MAR-2003 (revised)  
 DT 16-JUN-1998 (first entry)  
 XX  
 DE T-cell epitope peptide 72 from Japanese cypress pollen antigen Chao2.  
 XX  
 KW Japanese cypress pollen; antigen; T-cell epitope; Chao1; Chao2;  
 KW diagnosis; allergy; spring tree pollen disease; pollinosis.  
 XX  
 OS Chamaecyparis obtusa.  
 XX  
 PN WO9747648-A1.  
 XX  
 PD 18-DEC-1997.  
 XX  
 PF 12-JUN-1997; 97WO-JP002031.  
 XX  
 PR 14-JUN-1996; 96JP-00153527.  
 XX  
 PA (MEIP) MEIJI MILK PROD CO LTD.  
 XX  
 PI Kino K, Dairiri K;  
 XX  
 PI WPI; 1998-052242/05.  
 XX  
 DE T-cell epitope peptide portion of Japanese cypress pollen antigens Chao1  
 PT and Chao2 - used for diagnosis and treatment of spring tree pollen  
 PT disease.  
 XX  
 PS Claim 2; Page 48; 71pp; Japanese.  
 XX  
 CC The present sequence represents a T-cell epitope peptide from Japanese  
 CC cypress pollen antigen Chao2. The present invention describes peptides  
 CC which correspond to the T-cell epitope sites on Japanese cypress pollen  
 CC antigens Chao1 and Chao2. The peptides can be used as a reagent for the  
 CC diagnosis of allergy to Japanese cypress pollen, and as an antigen in the  
 CC treatment and prevention of spring tree pollen disease in which the  
 CC pollinosis involves reactivity to Japanese cypress pollen. (Updated on 25  
 CC -MAR-2003 to correct PI field.) (Updated on 27-AUG-2003 to correct OS  
 CC field.)  
 XX  
 SQ Sequence 20 AA;  
 Query Match 33.3%; Score 18; DB 2; Length 20;  
 Best Local Similarity 28.6%; Pred. No. 8.5e+03;  
 Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
 QY 3 CXXXXXC 9  
 DB 7 CTSASAC 13  
 RESULT 20  
 ADM10961  
 ID ADM10961 standard; peptide; 25 AA.  
 XX  
 AC ADM10961;  
 XX  
 DT 20-MAY-2004 (first entry)  
 XX  
 DE Human ovary-specific O1034C/O591S peptide #2.  
 XX  
 KW ovarian; tumour protein; cancer; T cell; CD4+ cell; CD8+ cell;  
 KW cytostatic; gene therapy; human; ds.  
 XX  
 OS Homo sapiens.



PR 21-SEP-2000; 2000US-0234687P.  
 PR 27-SEP-2000; 2000US-0236359P.  
 PR 04-OCT-2000; 2000GB-00024263.  
 XX  
 PA (MOLE-) MOLECULAR DYNAMICS INC.  
 XX Penn SG, Hanzel DK, Chen W, Rank DR;  
 XX WPI; 2001-483447/52.  
 XX  
 PT Human genome-derived single exon nucleic acid probes useful for analyzing  
 PT gene expression in human fetal liver.  
 XX  
 XX Claim 27; SEQ ID NO 34938; 639pp + Sequence Listing; English.  
 CC The invention relates to a single exon nucleic acid probe for measuring  
 CC human gene expression in a sample derived from human foetal liver. The  
 CC single exon nucleic acid probes may be used for predicting, measuring and  
 CC displaying gene expression in samples derived from human fetal liver. The  
 CC present sequence is a peptide encoded by a single exon nucleic acid probe  
 CC of the invention. Note: The sequence data for this patent did not form  
 CC part of the printed specification, but was obtained in electronic format  
 CC directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 XX Sequence 28 AA;  
 SQ  
 Query Match 33.3%; Score 18; DB 4; Length 28;  
 Best Local Similarity 28.6%; Pred. No. 1e+04;  
 Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
 QY 3 CXXXXXC 9  
 DB 3 CASSAC 9  
 RESULT 23  
 ID AAM36111 standard; protein; 28 AA.  
 XX  
 AC AAM36111;  
 DT 17-OCT-2001 (first entry)  
 XX  
 DE Peptide #10148 encoded by probe for measuring placental gene expression.  
 XX  
 KW Probe; microarray; human; placenta; antenatal diagnosis;  
 KW genetic disorder.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200157272-A2.  
 XX  
 PD 09-AUG-2001.  
 XX  
 PF 30-JAN-2001; 2001WO-US000663.  
 XX  
 PR 04-FEB-2000; 2000US-0180312P.  
 PR 26-MAY-2000; 2000US-0207456P.  
 PR 30-JUN-2000; 2000US-00608408.  
 PR 03-AUG-2000; 2000US-00632366.  
 PR 21-SEP-2000; 2000US-0234687P.  
 PR 04-OCT-2000; 2000GB-00024263.  
 XX  
 PA (MOLE-) MOLECULAR DYNAMICS INC.  
 XX Penn SG, Hanzel DK, Chen W, Rank DR;  
 XX WPI; 2001-488997/53.  
 XX  
 PT Human genome-derived single exon nucleic acid probes useful for analyzing  
 PT gene expression in human placenta.  
 XX

PS Claim 27; SEQ ID NO 36380; 654pp; English.  
 XX  
 CC The present invention relates to single exon nucleic acid probes (SENP:  
 CC see AAI31315-AAI57546). The present sequence is a peptide encoded by one  
 CC such probe. The probes are useful for producing a microarray for  
 CC predicting, measuring and displaying gene expression in samples derived  
 CC from human placenta. The probes are useful for antenatal diagnosis of  
 CC human genetic disorders  
 XX  
 XX Sequence 28 AA;  
 SQ  
 Query Match 33.3%; Score 18; DB 4; Length 28;  
 Best Local Similarity 28.6%; Pred. No. 1e+04;  
 Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
 QY 3 CXXXXXC 9  
 DB 3 CASSAC 9  
 RESULT 24  
 ID ABB25806 standard; protein; 28 AA.  
 XX  
 AC ABB25806;  
 XX  
 DT 23-JAN-2002 (first entry)  
 XX  
 DE Protein #7805 encoded by probe for measuring heart cell gene expression.  
 XX  
 KW Human; gene expression; heart; microarray; vascular system;  
 KW cardiovascular disease; hypertension; cardiac arrhythmia;  
 KW congenital heart disease.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200157274-A2.  
 XX  
 PD 09-AUG-2001.  
 XX  
 PF 30-JAN-2001; 2001WO-US000666.  
 XX  
 PR 04-FEB-2000; 2000US-0180312P.  
 PR 26-MAY-2000; 2000US-0207456P.  
 PR 30-JUN-2000; 2000US-00608408.  
 PR 03-AUG-2000; 2000US-00632366.  
 PR 21-SEP-2000; 2000US-0234687P.  
 PR 27-SEP-2000; 2000US-0236359P.  
 PR 04-OCT-2000; 2000GB-00024263.  
 XX  
 PA (MOLE-) MOLECULAR DYNAMICS INC.  
 XX Penn SG, Hanzel DK, Chen W, Rank DR;  
 XX WPI; 2001-488997/53.  
 XX  
 PT Single exon nucleic acid probes for analyzing gene expression in human  
 PT hearts.  
 XX  
 XX Claim 15; SEQ ID NO 27576; 530pp; English.  
 CC  
 CC The present invention relates to single exon nucleic acid probes for  
 CC measuring human gene expression in a sample derived from human heart (see  
 CC ABA21535-ABA41305). The present sequence is a protein encoded by one such  
 CC probe. The probes may be used for predicting, measuring and displaying  
 CC gene expression in samples derived from the human heart via microarrays.  
 CC By measuring gene expression, the probes are useful for predicting,  
 CC diagnosing, grading, staging, monitoring and prognosing diseases of the  
 CC human heart and vascular system e.g. cardiovascular disease,  
 CC hypertension, cardiac arrhythmias and congenital heart disease. Note: The  
 CC sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences

```
XX SQ Sequence 28 AA;
Query Match 33.3%; Score 18; DB 4; Length 28;
Best Local Similarity 28.6%; Pred. NO. 1e+04;
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 3 CXXXXXC 9
Db 3 CASSSAC 9

RESULT 25
AAM76003
ID AAM76003 standard; protein; 28 AA.
XX AC
XX DT 06-NOV-2001 (first entry)
XX DE Human bone marrow expressed probe encoded protein SEQ ID NO: 36309.
XX DE Human; bone marrow expressed exon; gene expression analysis; probe;
XX KW microarray; cancer; leukaemia; lymphoma; myeloma.
XX KW Homo sapiens.
XX OS
XX PN WO200157276-A2.
XX PD 09-AUG-2001.
XX PF 30-JAN-2001; 2001WO-US000668.
XX PR 04-FEB-2000; 2000US-0180312P.
XX PR 26-MAY-2000; 2000US-0207456P.
XX PR 30-JUN-2000; 2000US-00608408.
XX PR 03-AUG-2000; 2000US-00632366.
XX PR 21-SEP-2000; 2000US-0234687P.
XX PR 27-SEP-2000; 2000US-0236359P.
XX PR 04-OCT-2000; 2000GB-00024263.
XX PA (MOLE-) MOLECULAR DYNAMICS INC.
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-488900/53.
XX DR Human genome-derived single exon nucleic acid probes useful for analyzing
XX gene expression in human bone marrow.
XX PS Example 4; SEQ ID NO 36309; 658pp + Sequence Listing; English.
XX CC The present invention provides a number of single exon nucleic acid
XX CC probes which are derived from genomic sequences expressed in the human
XX CC bone marrow. They can be used to measure gene expression in bone marrow
XX CC samples, which may enable the improved diagnosis and treatment of cancers
XX CC such as lymphoma, leukaemia and myeloma. The present sequence is a
XX CC protein encoded by one of the probes of the invention
XX SQ Sequence 28 AA;
Query Match 33.3%; Score 18; DB 4; Length 28;
Best Local Similarity 28.6%; Pred. NO. 1e+04;
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 3 CXXXXXC 9
Db 3 CASSSAC 9

RESULT 26
AAM63189
ID AAM63189 standard; protein; 28 AA.
XX AC
XX DT 05-NOV-2001 (first entry)
XX DE Human brain expressed single exon probe encoded protein SEQ ID NO: 35294.
XX DE Human; brain expressed exon; gene expression analysis; probe; microarray;
XX KW Alzheimer's disease; multiple sclerosis; schizophrenia; epilepsy; cancer.
XX KW Homo sapiens.
XX OS
XX PN WO200157275-A2.
XX PD 09-AUG-2001.
XX PF 30-JAN-2001; 2001WO-US000667.
XX PR 04-FEB-2000; 2000US-0180312P.
XX PR 26-MAY-2000; 2000US-0207456P.
XX PR 30-JUN-2000; 2000US-00608408.
XX PR 03-AUG-2000; 2000US-00632366.
XX PR 21-SEP-2000; 2000US-0234687P.
XX PR 27-SEP-2000; 2000US-0236359P.
XX PR 04-OCT-2000; 2000GB-00024263.
XX PA (MOLE-) MOLECULAR DYNAMICS INC.
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-483446/52.
XX DR Single exon nucleic acid probes for analyzing gene expression in human
XX PT brains.
XX PS Example 4; SEQ ID NO 35294; 650pp + Sequence Listing; English.
XX CC The present invention provides a number of single exon nucleic acid
XX CC probes which are derived from genomic sequences expressed in the human
XX CC brain. They can be used to measure gene expression in brain cell samples,
XX CC which may enable the diagnosis and improved treatment of nervous system
XX CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia, of
XX CC epilepsy and cancers. The present sequence is a protein encoded by one of
XX CC the probes of the invention
XX SQ Sequence 28 AA;
Query Match 33.3%; Score 18; DB 4; Length 28;
Best Local Similarity 28.6%; Pred. NO. 1e+04;
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 3 CXXXXXC 9
Db 3 CASSSAC 9

RESULT 27
ABG57729
ID ABG57729 standard; peptide; 28 AA.
XX AC
XX DT 25-FEB-2003 (first entry)
XX DE Human liver peptide, SEQ ID NO 36377.
XX DE Human; liver; cirrhosis; hyperlipoproteinaemia; hyperlipidaemia;
XX KW hypercholesterolaemia; coronary heart disease.
XX KW Homo sapiens.
XX OS
XX PN WO200157273-A2.
XX SQ
```



RESULT 29  
ABO57875  
ID ABO57875 standard; protein; 39 AA.  
XX  
AC ABO57875;  
XX  
DT 29-JUL-2004 (first entry)  
XX  
DE Human genome derived single exon protein #4109.  
XX  
KW Human; gene expression; single exon probe; microarray;  
KW alternative splicing event; genomic alteration.  
XX  
OS Homo sapiens.  
XX  
PN US2003194704-A1.  
XX  
PD 16-OCT-2003.  
XX  
PF 03-APR-2002; 2002US-00029386.  
XX  
PR 03-APR-2002; 2002US-00029386.  
XX  
PA (PENN/) PENN S G.  
PA (RANK/) RANK D R.  
PA (HANZ/) HANZEL D K.  
XX  
PI Penn SG, Rank DR, Hanzel DK;  
XX  
DR WPI; 2004-119264/12.  
XX  
PT New human genome-derived single exon nucleic acid probes useful for human  
PT gene expression analysis, for identifying or characterizing alternative  
PT splicing events, for assessing genomic alterations or as tools for  
PT surveying tissues.  
XX  
PS Claim 45; SEQ ID NO 31509; 80pp; English.  
XX  
CC The invention relates to a nucleic acid probe for measuring human gene  
CC expression, comprising any of the 27,400 fully defined nucleotide  
CC sequences in the specification, or their complements or fragments, and  
CC encoding at least 8 amino acids of any of the 6888 amino acid sequences  
CC fully defined in the specification. The probe is a single exon probe that  
CC hybridizes under high stringency conditions to a nucleic acid molecule  
CC expressed in human cells or tissues. Also included are a spatially-  
CC addressable set of single exon nucleic acid probes for measuring human  
CC gene expression (comprising a plurality of single exon nucleic acid  
CC probes cited above, where each of the plurality of probes is separately  
CC and addressably isolatable or amplifiable from the plurality), a single  
CC exon microarray for measuring human gene expression, a method of  
CC measuring human gene expression, a vector comprising the single exon  
CC probe cited above, an ORF-encoded peptide comprising at least 8  
CC contiguous amino acids of any of the above-mentioned amino acid  
CC sequences (optionally with conservative amino acid substitutions), an  
CC isolated antibody that binds specifically to a peptide cited above,  
CC methods of selling and/or licensing single exon probes or microarrays to  
CC a customer desiring to measure gene expression, a method of providing  
CC human gene expression data by subscription, and a computer-readable  
CC storage medium which contains a database having a plurality of records  
CC (each record including data on the expression of a single exon probe  
CC cited above. The probe, methods and apparatus are useful in gene  
CC expression analysis. The probes may be used as tools for surveying  
CC tissues to detect the presence of expressed messages that contain their  
CC specific exon, or in constructing genome-derived single exon microarrays.  
CC In addition, the probes are used in identifying and characterising  
CC alternative splicing events, in detecting and characterising gross  
CC alterations in the genomic locus that includes their exon, in assessing  
CC smaller genomic alterations, in priming the synthesis of nucleic acids,  
CC or in expressing the ORF-encoded peptide. The present sequence is a human  
CC single exon probe protein of the invention. Note: The sequence data for  
CC this patent did not form part of the printed specification, but was  
CC obtained in electronic format directly from USPTO at  
CC seqdata.uspto.gov/sequence.html?DocID=20030194704

XX  
SQ Sequence 39 AA;  
Query Match 33.3%; Score 18; DB 8; Length 39;  
Best Local Similarity 28.6%; Pred. NO. 1.2e+04;  
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
QY 3 CXXXXXC 9  
DB 19 CSATSTC 25  
RESULT 30  
ABO5767  
ID ABO5767 standard; protein; 57 AA.  
XX  
AC ABO5767;  
XX  
DT 13-FEB-2002 (first entry)  
XX  
DE Novel human diagnostic protein #5758.  
XX  
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;  
KW food supplement; medical imaging; diagnostic; genetic disorder.  
XX  
OS Homo sapiens.  
XX  
PN WO200175067-A2.  
XX  
PD 11-OCT-2001.  
XX  
PF 30-MAR-2001; 2001WO-US008631.  
XX  
PR 31-MAR-2000; 2000US-00540217.  
PR 23-AUG-2000; 2000US-00649167.  
XX  
PA (HYSE-) HYSEQ INC.  
XX  
PI Drmanac RT, Liu C, Tang YT;  
XX  
DR WPI; 2001-639362/73.  
DR N-PSDB; AAS69954.  
XX  
PT New isolated polynucleotide and encoded polypeptides, useful in  
PT diagnostics, forensics, gene mapping, identification of mutations  
PT responsible for genetic disorders or other traits and to assess  
PT biodiversity.  
XX  
PS Claim 20; SEQ ID NO 36126; 103pp; English.  
XX  
CC The invention relates to isolated polynucleotide (I) and polypeptide (II)  
CC sequences. (I) is useful as hybridisation probes, polymerase chain  
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,  
CC and in recombinant production of (II). The polynucleotides are also used  
CC in diagnostics as expressed sequence tags for identifying expressed  
CC genes. (I) is useful in gene therapy techniques to restore normal  
CC activity of (II) or to treat disease states involving (II). (II) is  
CC useful for generating antibodies against it, detecting or quantitating a  
CC polypeptide in tissue, as molecular weight markers and as a food  
CC supplement. (II) and its binding partners are useful in medical imaging  
CC of sites expressing (II). (I) and (II) are useful for treating disorders  
CC involving aberrant protein expression or biological activity. The  
CC polypeptide and polynucleotide sequences have applications in  
CC diagnostics, forensics, gene mapping, identification of mutations  
CC responsible for genetic disorders or other traits to assess biodiversity  
CC and to produce other types of data and products dependent on DNA and  
CC amino acid sequences. ABO00010-ABG30377 represent novel human diagnostic  
CC amino acid sequences of the invention. Note: The sequence data for this  
CC patent did not appear in the printed specification, but was obtained in  
CC electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 57 AA;

Query Match 33.3%; Score 18; DB 4; Length 57;  
Best Local Similarity 28.6%; Pred. No. 1.4e+04;  
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 3 CXXXXXC 9  
Db 46 CSTSASC 52

Search completed: April 27, 2005, 16:44:29  
Job time : 172 secs

This Page Blank (usp1c)



Result No.	Score	Query		DB	ID	Description
		Match	Length			
1	19	35.2	103	2	S73440	hypothetical prote
2	18	33.3	121	1	RWBEM3	surface glycoprote
3	18	33.3	135	2	A45835	ly6 homolog RK10 p
4	18	33.3	157	2	S09805	hypothetical prote
5	18	33.3	295	2	T00399	probable AP2 domai
6	18	33.3	381	2	B97542	hypothetical prote
7	18	33.3	381	2	AD2761	rare lipoprotein A
8	18	33.3	404	2	JC5784	adrenomedullin rec
9	18	33.3	442	2	T21496	hypothetical prote
10	18	33.3	492	2	S49147	ERF-2 protein - hu
11	18	33.3	507	2	JC7366	Jun a 2 protein -
12	18	33.3	514	2	S48730	Cry j II protein -
13	18	33.3	514	2	JC2498	second major aller
14	18	33.3	514	2	JC7100	polygalacturonase
15	18	33.3	525	2	S53194	cdc25A - rat
16	18	33.3	584	1	C8HUA	complement C8 alph
17	18	33.3	675	2	D48492	kpsc protein - Esc
18	18	33.3	746	2	T24978	hypothetical prote
19	18	33.3	1817	2	T34249	hypothetical prote
20	17	31.5	25	1	SMWR	metallothionein -
21	17	31.5	26	1	SMNC	metallothionein -
22	17	31.5	26	2	S55029	CAP3 protein - ant
23	17	31.5	27	2	S42359	ferredoxin [3Fe-4S
24	17	31.5	45	1	FFVZ	fulvicoin C - Myxoc
25	17	31.5	45	2	AC3539	hypothetical prote
26	17	31.5	61	1	SMAY1C	metallothionein I
27	17	31.5	61	2	S00810	metallothionein Ic
28	17	31.5	66	2	S59621	metallothionein is
29	17	31.5	71	2	H95299	hypothetical prote

Wed Apr 27 17:41:54 2005

103 17 31.5 326 2 T27655  
104 17 31.5 330 2 T39511  
105 17 31.5 333 2 T32058  
106 17 31.5 335 2 A39743  
107 17 31.5 336 2 B71366  
108 17 31.5 336 2 C59811  
109 17 31.5 339 2 A95937  
110 17 31.5 343 2 A37377  
111 17 31.5 346 2 D42087  
112 17 31.5 346 2 S64693  
113 17 31.5 348 2 T40152  
114 17 31.5 352 2 T42744  
115 17 31.5 354 2 B97003  
116 17 31.5 356 2 T19520  
117 17 31.5 360 2 D84646  
118 17 31.5 360 2 S69063  
119 17 31.5 362 2 I37458  
120 17 31.5 362 2 T38428  
121 17 31.5 370 2 E70390  
122 17 31.5 371 2 T27643  
123 17 31.5 374 2 JC2123  
124 17 31.5 374 2 JC2124  
125 17 31.5 389 2 T34766  
126 17 31.5 392 2 A60777  
127 17 31.5 393 2 D70858  
128 17 31.5 396 2 T45399  
129 17 31.5 400 2 E88206  
130 17 31.5 403 2 E87710  
131 17 31.5 404 2 AH0803  
132 17 31.5 404 2 JS0073  
133 17 31.5 404 2 C97154  
134 17 31.5 405 2 C83443  
135 17 31.5 406 1 SYEAL1  
136 17 31.5 406 2 G84940  
137 17 31.5 406 2 G91029  
138 17 31.5 406 2 H85873  
139 17 31.5 406 2 A64128  
140 17 31.5 407 2 AE0336  
141 17 31.5 407 2 AH2594  
142 17 31.5 410 2 F87122  
143 17 31.5 412 1 KRSHL1  
144 17 31.5 416 2 A45510  
145 17 31.5 416 2 A61404  
146 17 31.5 416 2 S60034  
147 17 31.5 416 2 A45559  
148 17 31.5 433 2 T02655  
149 17 31.5 438 1 S70602  
150 17 31.5 438 2 D88206

ALIGNMENTS

RESULT 1  
S73440  
Hypothetical protein B01\_orf103b - Mycoplasma pneumoniae (strain ATCC 29342)  
C:Species: Mycoplasma pneumoniae  
A:Variety: ATCC 29342  
C>Date: 26-Feb-1997 #sequence\_revision 25-Apr-1997 #text\_change 09-Jul-2004  
R:Himmelsreich, R.; Hilbert, H.; Plagens, H.; Pirkel, E.; Li, B.C.; Herrmann, R.  
Nucleic Acids Res. 24, 4420-4449, 1996  
A:Title: Complete sequence analysis of the genome of the bacterium Mycoplasma pneumoniae  
A:Reference number: S73327, MUID:97105885, PMID:8948633  
A:Accession: S73440  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-103 <HIM>  
A:Cross-references: UNIPROT:P75074; EMBL:AE000014; GB:U00089; NID:g1673770; PIDN:AA89578  
A>Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1996  
C:Genetics:  
A:Genetic code: SGC3

Query Match 35.2%; Score 19; DB 2; Length 103;  
Best Local Similarity 13.6%; Pred. No. 2.4e+02;  
Matches 3; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

QY 3 CXXXXXXCXXXXXXXXXXXXXXH 24  
DB 53 CSSISFCSLASSARLRYSSSH 74

RESULT 2

RWBEM3  
surface glycoprotein CD59 precursor homolog - saimirine herpesvirus 1 (strain 11)  
C:Species: saimirine herpesvirus 1  
A:Note: host Saimiri sciureus (common squirrel monkey)  
C>Date: 31-Dec-1992 #sequence\_revision 31-Dec-1992 #text\_change 22-Jun-1999  
C:Accession: A43384; F36807  
R:Albrecht, J.C.; Nicholas, J.; Cameron, K.R.; Newman, C.; Fleckenstein, B.; Hones, R.W.  
Virology 190, 527-530, 1992  
A:Title: Herpesvirus saimiri has a gene specifying a homologue of the cellular membrane  
A:Reference number: A43384; MUID:92410640; PMID:1382344  
A:Accession: A43384  
A:Molecule type: DNA  
A:Residues: 1-121 <ALB>  
A:Cross-references: GB:S44811; GB:X64273; NID:G60318; PIDN:CAA45565.1; PID:G60319  
C:Genetics:  
A:Gene: 15  
C:Superfamily: Ly-6 antigen; Ly-6 homology  
C:Keywords: glycoprotein; lipoprotein; phosphatidylinositol linkage  
F:1-19/Domain: signal sequence #status predicted <SIG>  
F:20-96/Product: surface glycoprotein #status predicted <SGP>  
F:20-96/Domain: Ly-6 homology <LY6>  
F:97-121/Domain: carboxyl-terminal propeptide #status predicted <CPP>  
F:24/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 33.3%; Score 18; DB 1; Length 121;  
Best Local Similarity 28.6%; Pred. No. 4.8e+02;  
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 CXXXXXC 9  
DB 32 CTTSTSC 38

RESULT 3

A45835  
Ly6 homolog RK10 precursor - rat  
C:Species: Rattus norvegicus (Norway rat)  
C>Date: 03-Mar-1994 #sequence\_revision 03-Mar-1994 #text\_change 09-Jul-2004  
C:Accession: A45835; B45835  
R:Friedman, S.; Palfrey, R.G.E.; Sirlin, S.; Haemmerling, U.  
Immunogenetics 31, 104-111, 1990  
A:Title: Analysis of three distinct Ly6-A-related cDNA sequences isolated from rat kidney  
A:Reference number: A45835; MUID:90152758; PMID:2154400  
A:Accession: A45835  
A:Molecule type: mRNA  
A:Residues: 1-135 <FRI>  
A:Cross-references: UNIPROT:Q63317; GB:M30689; NID:G205247; PIDN:AAA41546.1; PID:G205248  
A:Experimental source: clone RK10  
A:Accession: B45835  
A:Molecule type: mRNA  
A:Residues: 25-54, 'A', 56-62, 'M', 64-67, 'Q', 69-71, 'DHI', 75-77, 'V', 79-80, 'T', 82-85, 'T', 87, '  
A:Cross-references: GB:M30692; NID:G205245; PIDN:AAA41545.1; PID:G205246  
A:Experimental source: clone RK6  
C:Superfamily: Ly-6 antigen; Ly-6 homology  
C:Keywords: blocked carboxyl end; glycoprotein; lipoprotein; phosphatidylinositol linkage  
F:1-26/Domain: signal sequence #status predicted <SIG>  
F:106/Modified site: GPI-anchor ethanolamine amidated carboxyl end (Asn) (in mature form)

Query Match 33.3%; Score 18; DB 2; Length 135;  
Best Local Similarity 28.6%; Pred. No. 5e+02;  
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 CXXXXXC 9

Db	41 CSSTATC 47	
RESULT 4		
S09805	hypothetical protein UL42 - human cytomegalovirus (strain AD169)	
C:Species:	human cytomegalovirus, human herpesvirus 5	
A:Note:	host Homo sapiens (man)	
C:Date:	07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 09-Jul-2004	
C:Accession:	S09805	
R:Chee, M.S.; Bankier, A.T.; Beck, S.; Bohni, R.; Brown, C.M.; Cerny, R.; Horsnell, T.; M.; Barrall, B.G.		
Curr. Top. Microbiol.	Immunol. 154, 125-169, 1990	
A:Title:	Analysis of the protein-coding content of the sequence of human cytomegalovirus	
A:Reference number:	S09749; MUID:90269039; PMID:2161319	
A:Accession:	S09805	
A:Status:	nucleic acid sequence not shown; translation not shown	
A:Molecule type:	DNA	
A:Residues:	1-157 <CHE>	
A:Cross-references:	UNIPROT:P16815; EMBL:X17403; NID:G59591; PIDN:CAA35401.1; PID:e27241	
A:Note:	this sequence was submitted to the EMBL Data Library, December 1989	
C:Superfamily:	human cytomegalovirus hypothetical protein UL42	
C:Keywords:	Glycoprotein; transmembrane protein	
F:117-143/Domain:	transmembrane #status predicted <TM>	
F:147/Binding site:	carbohydrate (Asn) (covalent) #status predicted	
Query Match	33.3%; Score 18; DB 2; Length 157;	
Best Local Similarity	28.6%; Pred. No. 5.4e+02;	
Matches	2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;	
QY	3 CXXXXXC 9	
Db	76 CAATSSC 82	
RESULT 5		
T00399	probable AP2 domain transcription factor [imported] - Arabidopsis thaliana	
N:Alternate names:	transcription factor TINY homolog T13E15.5	
C:Species:	Arabidopsis thaliana (mouse-ear cress)	
C:Date:	01-Feb-1999 #sequence_revision 01-Feb-1999 #text_change 09-Jul-2004	
C:Accession:	T00399; F84884	
R:Rounsley, S.D.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes, S.M.; Maso		
submitted to the EMBL Data Library, July 1997		
A:Description:	Arabidopsis thaliana chromosome II BAC T13E15 genomic sequence.	
A:Reference number:	Z14146	
A:Accession:	T00399	
A:Status:	translated from GB/EMBL/DBDJ	
A:Molecule type:	DNA	
A:Residues:	1-295 <ROU>	
A:Cross-references:	UNIPROT:Q22158; EMBL:AC002388; NID:G3420042; PID:G2344890	
A:Experimental source:	cultivar Columbia	
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.B.; Umayam, L.; Tallon, L.		
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J		
Nature 402, 761-768, 1999		
A:Title:	Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.	
A:Reference number:	A84420; MUID:20083487; PMID:10617197	
A:Accession:	F84884	
A:Status:	preliminary	
A:Molecule type:	DNA	
A:Residues:	1-295 <STO>	
A:Cross-references:	GB:AE002093; NID:G4895256; PIDN:RAD32841.1; GSPDB:GN00139	
C:Genetics:		
A:Gene:	T13E15.5; At2g44940	
A:Map position:	2	
Query Match	33.3%; Score 18; DB 2; Length 295;	
Best Local Similarity	28.6%; Pred. No. 7.1e+02;	
Matches	2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;	

QY	3 CXXXXXC 9	
Db	233 CASTSAC 239	
RESULT 6		
B97542	hypothetical protein AGR_C 2765 [imported] - Agrobacterium tumefaciens (strain C58, Cer:	
C:Species:	Agrobacterium tumefaciens	
C:Date:	30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 09-Jul-2004	
C:Accession:	B97542	
R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman		
A.; Liu, F.; Wollam, C.; Allinger, M.; Dougherty, D.; Scott, C.; Leppas, C.; Markelz, B.		
Science 294, 2323-2328, 2001		
A:Title:	Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tu	
A:Reference number:	A97359; MUID:21608551; PMID:11743194	
A:Accession:	B97542	
A:Status:	preliminary	
A:Molecule type:	DNA	
A:Residues:	1-381 <KUR>	
A:Cross-references:	UNIPROT:Q8UF98; GB:AE007869; PIDN:AAK87291.1; PID:gl5156585; GSPDB:	
C:Genetics:		
A:Gene:	AGR_C 2765	
A:Map position:	circular chromosome	
C:Superfamily:	rplA lipoprotein	
Query Match	33.3%; Score 18; DB 2; Length 381;	
Best Local Similarity	28.6%; Pred. No. 8e+02;	
Matches	2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;	
QY	3 CXXXXXC 9	
Db	44 CAATASC 50	
RESULT 7		
AD2761	rare lipoprotein A [imported] - Agrobacterium tumefaciens (strain C58, Dupont)	
C:Species:	Agrobacterium tumefaciens	
C:Date:	11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 09-Jul-2004	
C:Accession:	AD2761	
R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, Y.		
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutyavin, T.; Levy, R.; Li, M.; McClell		
; Karp, P.; Romero, P.; Zhang, S.		
Science 294, 2317-2323, 2001		
A:Authors:	Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,	
ster, E.W.		
A:Title:	The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.	
A:Reference number:	AB2577; MUID:21608550; PMID:11743193	
A:Accession:	AD2761	
A:Status:	preliminary	
A:Molecule type:	DNA	
A:Residues:	1-381 <KUR>	
A:Cross-references:	UNIPROT:Q8UF98; GB:AE008688; PIDN:AAL42506.1; PID:gl7739925; GSPDB:	
A:Experimental source:	strain C58 (Dupont)	
C:Genetics:		
A:Gene:	rlpA	
A:Map position:	circular chromosome	
C:Superfamily:	rplA lipoprotein	
Query Match	33.3%; Score 18; DB 2; Length 381;	
Best Local Similarity	28.6%; Pred. No. 8e+02;	
Matches	2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;	
QY	3 CXXXXXC 9	
Db	44 CAATASC 50	
RESULT 8		
JC5784	adrenomedullin receptor - human	
C:Species:	Homo sapiens (man)	

C;Date: 24-Jan-1998 #sequence\_revision 13-Mar-1998 #text\_change 09-Jul-2004  
C;Accession: J05784  
R;Haenze, J.; Dittlich, K.; Doetsch, J.; Rascher, W.  
Biochem. Biophys. Res. Commun. 240, 183-188, 1997  
A;Title: Molecular cloning of a novel human receptor gene with homology to the rat adreno-  
A;Reference number: J05784; MUID:98042541; PMID:9367907  
A;Accession: J05784  
A;Molecule type: mRNA  
A;Residues: 1-404 <HAE>  
A;Cross-references: UNIPROT:O15218; GB:Y13583; NID:g2652933; PIDN:CAA73910.1; PID:g26529  
C;Superfamily: vertebrate rhodopsin  
C;Keywords: glycoprotein; receptor; transmembrane protein  
F;58-79/Domain: transmembrane #status predicted <TM1>  
F;91-113/Domain: transmembrane #status predicted <TM2>  
F;128-149/Domain: transmembrane #status predicted <TM3>  
F;171-193/Domain: transmembrane #status predicted <TM4>  
F;218-239/Domain: transmembrane #status predicted <TM5>  
F;260-280/Domain: transmembrane #status predicted <TM6>  
F;301-320/Domain: transmembrane #status predicted <TM7>  
F;28,37/Binding site: carboxylate (Asn) (covalent) #status predicted

Query Match 33.3%; Score 18; DB 2; Length 404;  
Best Local Similarity 28.6%; Pred. No. 8.2e+02;  
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 3 CXXXXXC 9  
Db 349 CASSSSC 355

RESULT 9  
T24196  
hypothetical protein R11H6.2 - Caenorhabditis elegans  
C;Species: Caenorhabditis elegans  
C;Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
C;Accession: T24196  
R;Bardill, S.  
submitted to the EMBL Data Library, March 1997  
A;Reference number: Z19852  
A;Accession: T24196  
A;Status: preliminary; translated from GB/EMBL/DBDJ  
A;Molecule type: DNA  
A;Residues: 1-442 <WIL>  
A;Cross-references: UNIPROT:O45719; EMBL:Z93386; PIDN:CAB07645.1; GSPDB:GN000023; CESP:R1  
A;Experimental source: clone R11H6  
C;Genetics:  
A;Gene: CESP:R11H6.2  
A;Map position: 5  
A;Introns: 13/3; 68/3; 154/1; 284/1; 356/1

Query Match 33.3%; Score 18; DB 2; Length 442;  
Best Local Similarity 28.6%; Pred. No. 8.6e+02;  
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 3 CXXXXXC 9  
Db 10 CAASAC 16

RESULT 10  
S49147  
ERF-2 protein - human  
C;Species: Homo sapiens (man)  
C;Date: 16-Feb-1995 #sequence\_revision 12-May-1995 #text\_change 09-Jul-2004  
C;Accession: S49147; 137372  
R;Nie, X.F.; Maclean, R.N.; Kumar, V.; McKay, I.A.; Bustin, S.A.  
submitted to the EMBL Data Library, April 1994  
A;Reference number: S49147  
A;Accession: S49147  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-492 <NIE>  
A;Cross-references: UNIPROT:P47974

C;Date: 24-Jan-1998 #sequence\_revision 13-Mar-1998 #text\_change 09-Jul-2004  
C;Accession: J05784  
R;Haenze, J.; Dittlich, K.; Doetsch, J.; Rascher, W.  
Biochem. Biophys. Res. Commun. 240, 183-188, 1997  
A;Title: Molecular cloning of a novel human receptor gene with homology to the rat adreno-  
A;Reference number: J05784; MUID:98042541; PMID:9367907  
A;Accession: J05784  
A;Molecule type: mRNA  
A;Residues: 1-404 <HAE>  
A;Cross-references: UNIPROT:O15218; GB:Y13583; NID:g2652933; PIDN:CAA73910.1; PID:g26529  
C;Superfamily: vertebrate rhodopsin  
C;Keywords: glycoprotein; receptor; transmembrane protein  
F;58-79/Domain: transmembrane #status predicted <TM1>  
F;91-113/Domain: transmembrane #status predicted <TM2>  
F;128-149/Domain: transmembrane #status predicted <TM3>  
F;171-193/Domain: transmembrane #status predicted <TM4>  
F;218-239/Domain: transmembrane #status predicted <TM5>  
F;260-280/Domain: transmembrane #status predicted <TM6>  
F;301-320/Domain: transmembrane #status predicted <TM7>  
F;28,37/Binding site: carboxylate (Asn) (covalent) #status predicted

Query Match 33.3%; Score 18; DB 2; Length 404;  
Best Local Similarity 28.6%; Pred. No. 8.2e+02;  
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 3 CXXXXXC 9  
Db 349 CASSSSC 355

RESULT 9  
T24196  
hypothetical protein R11H6.2 - Caenorhabditis elegans  
C;Species: Caenorhabditis elegans  
C;Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
C;Accession: T24196  
R;Bardill, S.  
submitted to the EMBL Data Library, March 1997  
A;Reference number: Z19852  
A;Accession: T24196  
A;Status: preliminary; translated from GB/EMBL/DBDJ  
A;Molecule type: DNA  
A;Residues: 1-442 <WIL>  
A;Cross-references: UNIPROT:O45719; EMBL:Z93386; PIDN:CAB07645.1; GSPDB:GN000023; CESP:R1  
A;Experimental source: clone R11H6  
C;Genetics:  
A;Gene: CESP:R11H6.2  
A;Map position: 5  
A;Introns: 13/3; 68/3; 154/1; 284/1; 356/1

Query Match 33.3%; Score 18; DB 2; Length 442;  
Best Local Similarity 28.6%; Pred. No. 8.6e+02;  
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 3 CXXXXXC 9  
Db 10 CAASAC 16

RESULT 10  
S49147  
ERF-2 protein - human  
C;Species: Homo sapiens (man)  
C;Date: 16-Feb-1995 #sequence\_revision 12-May-1995 #text\_change 09-Jul-2004  
C;Accession: S49147; 137372  
R;Nie, X.F.; Maclean, R.N.; Kumar, V.; McKay, I.A.; Bustin, S.A.  
submitted to the EMBL Data Library, April 1994  
A;Reference number: S49147  
A;Accession: S49147  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-492 <NIE>  
A;Cross-references: UNIPROT:P47974

R;Nie, X.F.; Maclean, K.N.; Kumar, V.; McKay, I.A.; Bustin, S.A.  
Gene 152, 285-286, 1995  
A;Title: ERF-2, the human homologue of the murine Tisild early response gene.  
A;Reference number: I37372; MUID:95137407; PMID:7835719  
A;Accession: I37372  
A;Status: preliminary; translated from GB/EMBL/DBDJ  
A;Molecule type: mRNA  
A;Residues: 1-492 <RES>  
A;Cross-references: EMBL:X78992; NID:g509777; PIDN:CAA55592.1; PID:g509778  
C;Genetics:  
A;Gene: GDB:BRF2; ERF2  
A;Cross-references: GDB:511261  
A;Map position: I3q12.3-13q12.3

Query Match 33.3%; Score 18; DB 2; Length 492;  
Best Local Similarity 28.6%; Pred. No. 9e+02;  
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 3 CXXXXXC 9  
Db 293 CSSASSC 299

RESULT 11  
JC7366  
Jun a 2 protein - mountain cedar  
C;Species: Juniperus ashei (mountain cedar)  
C;Date: 17-Nov-2000 #sequence\_revision 17-Nov-2000 #text\_change 09-Jul-2004  
C;Accession: JC7366; PC7093  
R;Yokoyama, M.; Miyahara, M.; Shimizu, K.; Kino, K.; Tsunoo, H.  
Biochem. Biophys. Res. Commun. 275, 195-202, 2000  
A;Title: Purification, identification, and cDNA cloning of Jun a 2, the second major all-  
A;Reference number: JC7366  
A;Accession: JC7366  
A;Molecule type: mRNA  
A;Residues: 1-507 <YOK>  
A;Cross-references: UNIPROT:Q9FY19; GB:AJ404653  
A;Accession: PC7093  
A;Molecule type: protein  
A;Residues: 55-63 <Y02>  
C;Comment: This protein, a second major allergen of mountain cedar pollen, which is invo-  
o the polygalacturonase family.  
C;Keywords: glycoprotein; pollen

Query Match 33.3%; Score 18; DB 2; Length 507;  
Best Local Similarity 28.6%; Pred. No. 9.1e+02;  
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 3 CXXXXXC 9  
Db 348 CTSAAAC 354

RESULT 12  
S48730  
Cry j II protein - Japanese cedar  
C;Species: Cryptomeria japonica (Japanese cedar)  
C;Date: 07-May-1995 #sequence\_revision 21-Jul-1995 #text\_change 09-Jul-2004  
C;Accession: S48730  
R;Namba, M.; Kurose, M.; Torigoe, K.; Hino, K.; Taniguchi, Y.; Fukuda, S.; Usui, M.; Ku-  
FEBS Lett. 353, 124-128, 1994  
A;Title: Molecular cloning of the second major allergen, Cry j II, from Japanese cedar  
A;Reference number: S48730; MUID:95010777; PMID:7926035  
A;Accession: S48730  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-514 <NAM>  
A;Cross-references: UNIPROT:P43212; GB:D37765; NID:g577695; PIDN:BAA07021.1; PID:dl0075

Query Match 33.3%; Score 18; DB 2; Length 514;  
Best Local Similarity 28.6%; Pred. No. 9.2e+02;  
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```

QY      3 CXXXXXC 9
DB      347 CTSASAC 353

RESULT 13
JC2498
second major allergen Cry j II precursor - Japanese cedar
C:Species: Cryptomeria japonica (Japanese cedar)
C>Date: 16-Mar-1995 #sequence_revision 26-May-1995 #text_change 09-Jul-2004
C:Accession: JC2498; PC2346; A60147
R:Komiyama, N.; Sone, T.; Shimizu, K.; Morikubo, K.; Kino, K.
Biochem. Biophys. Res. Commun. 201, 1021-1028, 1994
A:Title: cDNA cloning and expression of Cry j II, the second major allergen of Japanese
A:Reference number: JC2498; MUID:94271186; PMID:8002372
A:Accession: JC2498
A:Molecule type: mRNA
A:Residues: 1-514 <KOM>
A:Cross-references: UNIPROT:P43212; DBJ:D29772; NID:G506857; PIDN:BAA06172.1; PID:G5068
A:Accession: PC2346
A:Molecule type: protein
A:Residues: 52-61 <KO2>
R:Sakaguchi, M.; Inouye, S.; Tani, M.; Ando, S.; Ueui, M.; Matuhasi, T.
Allergy 45, 309-312, 1990
A:Title: Identification of the second major allergen of Japanese cedar pollen.
A:Reference number: A60147; MUID:90342988; PMID:2382797
A:Accession: A60147
A:Molecule type: protein
A:Residues: 55-64 <SAK>
C:Keywords: glycoprotein; pollen
F:1-54/Domain: signal sequence #status predicted <SIG>
F:55-460/Product: second major allergen Cry j #status predicted <MAT>
F:429,460,472/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match      33.3%; Score 18; DB 2; Length 514;
Best Local Similarity 28.6%; Pred. No. 9.2e+02;
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY      3 CXXXXXC 9
DB      347 CTSASAC 353

RESULT 14
JC7100
polysaccharuronase Cha o 2 - Japanese cypress
C:Species: Chamaecyparis obtusa (Japanese cypress)
C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C:Accession: JC7100; PC7026
R:Mori, T.; Yokoyama, M.; Komiyama, N.; Okano, M.; Kino, K.
Biochem. Biophys. Res. Commun. 263, 166-171, 1999
A:Title: Purification, identification, and cDNA cloning of Cha o 2, the second major all
A:Reference number: JC7100; MUID:99417540; PMID:10486272
A:Accession: JC7100
A:Molecule type: mRNA
A:Residues: 1-514 <MOR>
A:Cross-references: UNIPROT:Q7M1E7
A:Accession: PC7026
A:Molecule type: protein
A:Residues: 51-62 <MO2>

Query Match      33.3%; Score 18; DB 2; Length 514;
Best Local Similarity 28.6%; Pred. No. 9.2e+02;
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY      3 CXXXXXC 9
DB      347 CTSASAC 353

RESULT 15
cdc25A - rat

```

```

C:Species: Rattus sp. (rat)
C>Date: 29-May-1998 #sequence_revision 29-May-1998 #text_change 22-Jun-1999
C:Accession: I53194
R:Jinno, S.; Suto, K.; Nagata, A.; Igarashi, M.; Kanaoka, Y.; Nojima, H.; Okayama, H.
EMBO J. 13, 1549-1556, 1994
A:Title: Cdc25A is a novel phosphatase functioning early in the cell cycle.
A:Reference number: I53194; MUID:94208523; PMID:8156993
A:Accession: I53194
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-525 <RES>
A:Cross-references: GB:D16236; NID:G484291; PIDN:BAA03761.1; PID:G1008041
C:Superfamily: human protein-tyrosine-phosphatase cdc25A; cdc25-type protein-tyrosine-ph
F:314-507/Domain: cdc25-type protein-tyrosine-phosphatase homology <PTP>

Query Match      33.3%; Score 18; DB 2; Length 525;
Best Local Similarity 28.6%; Pred. No. 9.2e+02;
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY      3 CXXXXXC 9
DB      266 CSSTSSC 272

RESULT 16
CB8UA
complement C8 alpha chain precursor [validated] - human
C:Species: Homo sapiens (man)
C>Date: 28-Dec-1987 #sequence_revision 17-Nov-2000 #text_change 17-Nov-2000
C:Accession: I37213; A26704
R:Michelotti, G.A.; Snider, J.V.; Sodetz, J.M.
Hum. Genet. 95, 513-518, 1995
A:Title: Genomic organization of human complement protein C8 alpha and further examinatio
A:Reference number: I37213; MUID:95278905; PMID:7759071
A:Accession: I37213
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-584 <MIC>
R:Rao, A.G.; Howard, O.M.Z.; Ng, S.C.; Whitehead, A.S.; Colten, H.R.; Sodetz, J.M.
Biochemistry 26, 3556-3564, 1987
A:Title: Complementary DNA and derived amino acid sequence of the alpha subunit of human
A:Reference number: A26704; MUID:88000560; PMID:2820471
A:Accession: A26704
A:Molecule type: mRNA
A:Residues: 1-92,'Q',94-466,'CCGTQAWASGQ',480-574,'P',576-584 <RAO>
A:Note: part of the sequence was confirmed by protein sequencing
R:Hotsteenge, J.; Blommestein, M.; Hess, D.; Furmanek, A.; Miroshnichenko, O.
J. Biol. Chem. 274, 32786-32794, 1999
A:Title: The four terminal components of the complement system are C-mannosylated on mul
A:Reference number: A59362; MUID:20020247; PMID:10551839
A:Contents: annotation
A:Note: identification and location of C-mannosylation sites by mass-spectroscopy
C:Genetics:
A:Gene: GDB:C8A
A:Cross-references: GDB:119735; OMIM:120950
A:Map position: lp32-lp32
A:Introns: 26/2, 57/3, 106/1, 155/2, 218/3, 285/3, 366/1, 408/1, 460/3, 535/1
C:Complex: heterotrimer of C8 alpha chain (PIR:CB8UA), C8 beta chain (PIR:CB8UB), and C8
C:Function:
A:Description: combines with complement C5b-7 complex to polymerize complement component
A:Pathway: complement pathway
C:Superfamily: complement C9; EGF homology; LDL receptor ligand-binding repeat homology;
C:Keywords: complement pathway; cytolysis; glycoprotein; membrane attack complex; plasma
F:1-20/Domain: signal sequence #status predicted <SIG>
F:21-30/Domain: propeptide #status predicted <PRO>
F:31-584/Product: complement C8 alpha chain #status predicted <MPT>
F:37-91/Domain: thrombospondin type 1 repeat homology <THR1>
F:96-130/Domain: LDL receptor ligand-binding repeat homology <LDL>
F:497-528/Domain: EGF homology <EGF>
F:538-584/Domain: thrombospondin type 1 repeat homology <THR2>
F:43/Binding site: carbohydrate (Asn) (covalent) #status absent
F:44,542,545,548/Modified site: 2'-mannosyl-tryptophan (Trp) #status experimental

```

F;437/Binding site: carbohydrate (Asn) (covalent):#status predicted

Query Match 33.3%; Score 18; DB 1; Length 584;  
Best Local Similarity 28.6%; Pred. No. 9.7e+02;  
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 CXXXXXC 9  
DB 84 CSSSTTC 90

## RESULT 17

D48492  
kpsc protein - Escherichia coli  
C:Species: Escherichia coli  
C>Date: 07-Apr-1994 #sequence\_revision 07-Apr-1994 #text\_change 09-Jul-2004  
C:Accession: D48492; I67694; S36652  
R:Pazzani, C.; Rosenow, C.; Boulnois, G.J.; Bronner, D.; Jann, K.; Roberts, I.S.  
J. Bacteriol. 175, 5978-5983, 1993  
A:Title: Molecular analysis of region 1 of the Escherichia coli K5 antigen gene cluster:  
A:Reference number: A48492; MUID:93388530; PMID:8397187  
A:Accession: D48492  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-675 <PAZ>  
A:Cross-references: UNIPROT:P42217; EMBL:X74567; NID:G397404; PIDN:CAA52658.1; PID:G3974  
R:Rosenow, C.; Roberts, I.S.; Jann, K.  
FEMS Microbiol. Lett. 125, 159-164, 1995  
A:Title: Isolation from recombinant Escherichia coli and characterization of CMP-Kdo synthetase from Escherichia coli  
A:Reference number: I53591; MUID:95180691; PMID:7875563  
A:Accession: I67694  
A:Status: preliminary; translated from GB/EMBL/DBDJ  
A:Molecule type: DNA  
A:Residues: 1-8 <RES>  
A:Cross-references: GB:S76943; NID:G913365

Query Match 33.3%; Score 18; DB 2; Length 675;  
Best Local Similarity 28.6%; Pred. No. 1e+03;  
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 CXXXXXC 9  
DB 361 CTAASAC 367

## RESULT 18

T24978  
hypothetical protein T19B10.3 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C>Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
C:Accession: T24978  
R:Lennard, N.  
submitted to the EMBL Data Library, June 1996  
A:Reference number: Z19964  
A:Accession: T24978  
A:Status: preliminary; translated from GB/EMBL/DBDJ  
A:Molecule type: DNA  
A:Residues: 1-746 <WIL>  
A:Cross-references: UNIPROT:Q27526; EMBL:Z74043; PIDN:CAA98542.1; GSPDB:GN00023; CESP:T1  
A:Experimental source: clone T19B10  
C:Genetics:  
A:Gene: CESP:T19B10.3  
A:Map position: 5  
A:Introns: 130/2; 198/2; 227/1; 257/2; 324/1; 479/2; 532/3; 605/3

Query Match 33.3%; Score 18; DB 2; Length 746;  
Best Local Similarity 28.6%; Pred. No. 1.1e+03;  
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 CXXXXXC 9  
DB 727 CTSTSSC 733

## RESULT 19

T34249  
hypothetical protein F31D5.5 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C>Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 09-Jul-2004  
C:Accession: T34249  
R:Wilcox, L.  
submitted to the EMBL Data Library, June 1995  
A:Description: The sequence of C. elegans cosmid F31D5.  
A:Reference number: Z21494  
A:Accession: T34249  
A:Status: preliminary; translated from GB/EMBL/DBDJ  
A:Molecule type: DNA  
A:Residues: 1-1817 <WIL>  
A:Cross-references: UNIPROT:Q19311; EMBL:U28941; PIDN:AAC71101.1; GSPDB:GN00020; CESP:F31  
A:Experimental source: strain Bristol N2; clone F31D5  
C:Genetics:  
A:Gene: CESP:F31D5.5  
A:Map position: 2  
A:Introns: 22/2; 107/2; 199/2; 291/2; 384/2; 476/2; 566/2; 648/2; 728/2; 904/2; 1047/2;

Query Match 33.3%; Score 18; DB 2; Length 1817;  
Best Local Similarity 28.6%; Pred. No. 1.6e+03;  
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 CXXXXXC 9  
DB 448 CTTTTTC 454

## RESULT 20

SMNR  
metallothionein - cultivated mushroom  
C:Species: Agaricus bisporus (cultivated mushroom)  
C>Date: 17-Mar-1987 #sequence\_revision 17-Mar-1987 #text\_change 09-Jul-2004  
C:Accession: A03286  
R:Munger, K.; Lerch, K.  
Biochemistry 24, 6751-6756, 1985  
A:Title: Copper metallothionein from the fungus Agaricus bisporus: chemical and spectroscopic studies  
A:Reference number: A03286  
A:Accession: A03286  
A:Molecule type: protein  
A:Residues: 1-25 <MUN>  
A:Cross-references: UNIPROT:P04358  
A:Experimental source: mycelium; strain A-32  
C:Comment: In contrast to the vertebrate metallothioneins, the fungal proteins bind copper  
C:Comment: The absorptive, luminescent, and stereoptical properties of the copper MT are  
C:Superfamily: metallothionein  
C:Keywords: chelation; metal binding; metal-thiolate cluster

Query Match 31.5%; Score 17; DB 1; Length 25;  
Best Local Similarity 28.6%; Pred. No. 4.4e+02;  
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 CXXXXXC 9  
DB 5 CSGASSC 11

## RESULT 21

SMNC  
metallothionein - Neurospora crassa  
C:Species: Neurospora crassa  
C>Date: 31-May-1980 #sequence\_revision 31-Mar-1993 #text\_change 09-Jul-2004  
C:Accession: A24641; A03287  
R:Muenzer, K.; Germann, U.A.; Lerch, K.  
EMBO J. 4, 2665-2668, 1985  
A:Title: Isolation and structural organization of the Neurospora crassa copper metallothionein  
A:Reference number: A24641; MUID:86030247; PMID:2932331  
A:Accession: A24641  
A:Molecule type: DNA  
A:Residues: 1-26 <MUE>

A;Cross-references: UNIPROT:P02807; GB:X03009; NID:g2986; PIDN:CAA26793.1; PID:g2987  
 R;Lerch, K.  
 Nature 284, 368-370, 1980  
 A;Title: Copper metallothionein, a copper-binding protein from *Neurospora crassa*.  
 A;Reference number: A03287; MUID:80143244; PMID:6444697  
 A;Accession: A03287  
 A;Molecule type: protein  
 A;Residues: 2-26 <LER>  
 C;Genetics:  
 A;Introns: 18/1  
 C;Superfamily: metallothionein  
 C;Keywords: chelation; metal binding; metal-thiolate cluster  
 F;4,6,12,14,18,20,23/Binding site: transition metal ions (Cys) #status experimental

Query Match 31.5%; Score 17; DB 1; Length 26;  
 Best Local Similarity 28.6%; Pred. No. 4.5e+02;  
 Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 CXXXXXC 9  
 |  
 Db 6 CSGASSC 12

RESULT 22  
 S55029  
 CAP3 protein - anthracnose fungus (*Colletotrichum gloeosporioides*)  
 C;Species: *Glomerella cingulata*, *Colletotrichum gloeosporioides*  
 C;Date: 23-Aug-1995 #sequence\_revision 19-Oct-1995 #text\_change 09-Jul-2004  
 A;Accession: S55029  
 R;Hwang, C.S.; Kolattukudy, P.E.  
 Mol. Gen. Genet. 247, 282-294, 1995  
 A;Title: Isolation and characterization of genes expressed uniquely during appressorium  
 A;Reference number: S55029; MUID:95287848; PMID:7770033  
 A;Accession: S55029  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-26 <HWA>  
 A;Cross-references: UNIPROT:Q99334; EMBL:U18756; NID:g619921; PIDN:AAA77679.1; PID:g6199  
 C;Genetics:  
 A;Introns: 24/3  
 C;Superfamily: metallothionein

Query Match 31.5%; Score 17; DB 2; Length 26;  
 Best Local Similarity 28.6%; Pred. No. 4.5e+02;  
 Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 CXXXXXC 9  
 |  
 Db 6 CASTGTC 12

RESULT 23  
 S42359  
 ferredoxin [3Fe-4S] - *Giardia lamblia* (fragment)  
 C;Species: *Giardia lamblia*  
 C;Date: 07-Oct-1994 #sequence\_revision 01-Mar-1996 #text\_change 12-Jul-2004  
 A;Accession: S42359  
 R;Townson, S.M.; Hanson, G.R.; Upcroft, J.A.; Upcroft, P.  
 Eur. J. Biochem. 220, 439-446, 1994  
 A;Title: A purified ferredoxin from *Giardia duodenalis*.  
 A;Reference number: S42359; MUID:94170792; PMID:8125101  
 A;Accession: S42359  
 A;Molecule type: protein  
 A;Residues: 1-27 <TOW>  
 A;Note: the source is given as *Giardia duodenalis*  
 C;Keywords: 3Fe-4S; iron-sulfur protein; metalloprotein  
 F;2-24/Domain: ferredoxin 2[4Fe-4S] homology (fragment) <FER>  
 F;9,15/Binding site: 3Fe-4S cluster (Cys) (covalent) #status experimental

Query Match 31.5%; Score 17; DB 2; Length 27;  
 Best Local Similarity 28.6%; Pred. No. 4.6e+02;  
 Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 CXXXXXC 9  
 |  
 Db 9 CSAAGAC 15

RESULT 24  
 FFYZ  
 fulvocin C - *Myxococcus fulvus*  
 C;Species: *Myxococcus fulvus*  
 C;Date: 14-Nov-1983 #sequence\_revision 14-Nov-1983 #text\_change 09-Jul-2004  
 A;Accession: A01810  
 R;Tsai, H.; Hirsch, H.J.  
 Biochim. Biophys. Acta 667, 213-217, 1981  
 A;Title: The primary structure of fulvocin C from *Myxococcus fulvus*.  
 A;Reference number: A01810; MUID:81161107; PMID:6783114  
 A;Accession: A01810  
 A;Molecule type: protein  
 A;Residues: 1-45 <TSA>  
 A;Cross-references: UNIPROT:P01547  
 C;Comment: Fulvocin C is a bacteriocin.  
 C;Comment: *Myxococcus fulvus* is a myxobacterium; these obligate aerobes are unicellular,  
 C;Superfamily: fulvocin C  
 C;Keywords: antibacterial

Query Match 31.5%; Score 17; DB 1; Length 45;  
 Best Local Similarity 28.6%; Pred. No. 5.8e+02;  
 Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 CXXXXXC 9  
 |  
 Db 18 CTTGTAC 24

RESULT 25  
 AC3539  
 hypothetical protein BMII10237 [imported] - *Brucella melitensis* (strain 16M)  
 C;Species: *Brucella melitensis*  
 C;Date: 01-Feb-2002 #sequence\_revision 01-Feb-2002 #text\_change 09-Jul-2004  
 A;Accession: AC3539  
 R;DelVecchio, V.G.; Kaputal, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova,  
 Mazur, M.; Goltzman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letes  
 Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002  
 A;Title: The genome sequence of the facultative intracellular pathogen *Brucella melitensis*  
 A;Reference number: AD3252; PMID:11756688  
 A;Accession: AC3539  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-45 <KUR>  
 A;Cross-references: UNIPROT:Q8YDE0; GB:AE008918; PIDN:AAL53478.1; PID:g17984380; GSPDB:  
 A;Experimental source: strain 16M  
 C;Genetics:  
 A;Gene: BMII10237  
 A;Map position: II

Query Match 31.5%; Score 17; DB 2; Length 45;  
 Best Local Similarity 28.6%; Pred. No. 5.8e+02;  
 Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 CXXXXXC 9  
 |  
 Db 22 CATSSQC 28

RESULT 26  
 SMHYIC  
 metallothionein I - Chinese hamster  
 C;Species: *Cricetulus griseus* (Chinese hamster)  
 C;Date: 17-May-1985 #sequence\_revision 17-May-1985 #text\_change 09-Jul-2004  
 A;Accession: A03282; S13737; S50198  
 R;Griffith, B.B.; Walters, R.A.; Enger, M.D.; Hildebrand, C.E.; Griffith, J.K.  
 Nucleic Acids Res. 11, 901-910, 1983  
 A;Title: cDNA cloning and nucleotide sequence comparison of Chinese hamster metallothior  
 A;Reference number: A93501; MUID:83168910; PMID:6687636

A;Accession: A03282  
A;Molecule type: mRNA  
A;Residues: 1-61 <GR1>  
A;Cross-references: UNIPROT:P02804; GB:J00061; NID:G191135; PIDN:AAA36996.1; PID:G304522  
R;Grady, D.L.; Robinson, D.L.; Hildebrand, C.E.  
Nucleic Acids Res. 18, 7149, 1990  
A;Title: Genomic sequence of the Chinese hamster MT I gene.  
A;Reference number: S13737; MUID:91088310; PMID:2263484  
A;Accession: S13737  
A;Molecule type: DNA  
A;Residues: 1-61 <GRA>  
A;Cross-references: EMBL:X5064; NID:q49468; PIDN:CAA38897.1; PID:q49469  
R;Yamada, K.; Kato, H.; Kanda, N.; Fujii-Kuriyama, Y.; Utakoji, T.; Itoh, R.  
Biochim. Biophys. Acta 1219, 581-591, 1994  
A;Title: Sequence homology of Chinese hamster metallothionein genes I and II to those of  
A;Reference number: S50198; MUID:95035087; PMID:7948015  
A;Accession: S50198  
A;Molecule type: DNA  
A;Residues: 1-61 <YAM>  
A;Cross-references: GB:D10551; GB:D90509; NID:G633055; PIDN:BAA01408.1; PID:G633056  
C;Genetics:  
A;Gene: MT-I  
A;Introns: 10/1; 32/1  
A;Superfamily: metallothionein  
C;Keywords: metal binding  
F;5,7,13,15,19,21,24,26,29/Binding site: transition metal ions (Cys) #status predicted  
F;33,34,36,37,41,44,48,50,57,59,60/Binding site: transition metal ions (Cys) #status pre  
Query Match 31.5%; Score 17; DB 1; Length 61;  
Best Local Similarity 28.6%; Pred. No. 6.6e+02; Mismatches 5; Indels 0; Gaps 0;  
Matches 2; Conservative 0;  
QY 3 CXXXXXC 9  
DB 7 CSTGSTC 13  
RESULT 27  
S00810  
metallothionein Ic - sheep  
C;Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)  
C;Date: 28-Feb-1990 #sequence\_revision 28-Feb-1990 #text\_change 09-Jul-2004  
C;Accession: S00810  
R;Peterson, M.G.; Hannan, F.; Mercer, J.F.B.  
Eur. J. Biochem. 174, 417-424, 1988  
A;Title: The sheep metallothionein gene family. Structure, sequence and evolutionary rel  
A;Reference number: S00808; MUID:88254812; PMID:3383853  
A;Accession: S00810  
A;Molecule type: DNA  
A;Residues: 1-61 <PET>  
A;Cross-references: UNIPROT:P09578; EMBL:X07974; NID:G1340; PIDN:CAA30786.1; PID:G1341  
C;Genetics:  
A;Introns: 10/1; 32/1  
C;Superfamily: metallothionein  
Query Match 31.5%; Score 17; DB 2; Length 61;  
Best Local Similarity 28.6%; Pred. No. 6.6e+02; Mismatches 5; Indels 0; Gaps 0;  
Matches 2; Conservative 0;  
QY 3 CXXXXXC 9  
DB 7 CSTGSSC 13  
RESULT 28  
S59621  
metallothionein isoform b, cadmium-binding - Ariantha arbustorum (terrestrial snail)  
C;Species: Ariantha arbustorum  
C;Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 09-Jul-2004  
C;Accession: S59621  
R;Berger, B.; Hunziker, P.E.; Hauer, C.R.; Birchler, N.; Dallinger, R.  
Biochem. J. 311, 951-957, 1995  
A;Title: Mass spectrometry and amino acid sequencing of two cadmium-binding metallothion

A;Reference number: S59621; MUID:96067616; PMID:7487956  
A;Accession: S59621  
A;Molecule type: protein  
A;Residues: 1-66 <BER>  
A;Cross-references: UNIPROT:P55946  
C;Superfamily: metallothionein  
C;Keywords: acetylated amino end; chelation; metal binding; metal-thiolate cluster  
F;1/Modified site: acetylated amino end (Ser) #status experimental  
Query Match 31.5%; Score 17; DB 2; Length 66;  
Best Local Similarity 28.6%; Pred. No. 6.9e+02; Mismatches 5; Indels 0; Gaps 0;  
Matches 2; Conservative 0;  
QY 3 CXXXXXC 9  
DB 51 CTGAASC 57  
RESULT 29  
H95299  
hypothetical protein Sma0580 [imported] - Sinorhizobium meliloti (strain 1021) magaplaasm  
C;Species: Sinorhizobium meliloti  
C;Date: 24-Aug-2001 #sequence\_revision 24-Aug-2001 #text\_change 09-Jul-2004  
C;Accession: H95299  
R;Barnett, M.J.; Fisher, R.F.; Jones, T.; Komp, C.; Abola, A.P.; Barloy-Hubler, F.; Bows  
; Kalman, S.; Keating, D.H.; Palm, C.; Peck, M.C.; Surzycki, R.; Wells, D.H.; Yeh, K.C.  
Proc. Natl. Acad. Sci. U.S.A. 98, 9883-9888, 2001  
A;Title: Nucleotide sequence and predicted functions of the entire Sinorhizobium melilot  
A;Reference number: A95262; MUID:21396509; PMID:11481432  
A;Accession: H95299  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-71 <KUR>  
A;Cross-references: UNIPROT:Q92Z24; GB:AE006469; PIDN:AAK64962.1; PID:G14523387; GSPDB:G  
A;Experimental source: strain 1021, megaplasamid pSymA  
R;Galibert, F.; Finan, T.M.; Long, S.K.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler,  
pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;  
L.; Hyman, R.W.; Jones, T.  
Science 293, 668-672, 2001  
A;Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,  
hebaull, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.  
A;Title: The composite genome of the legume symbiont Sinorhizobium meliloti.  
A;Reference number: A96039; MUID:21368234; PMID:11474104  
A;Contents: annotation  
C;Genetics:  
A;Gene: Sma0580  
A;Genome: plasmid  
Query Match 31.5%; Score 17; DB 2; Length 71;  
Best Local Similarity 28.6%; Pred. No. 7.1e+02; Mismatches 5; Indels 0; Gaps 0;  
Matches 2; Conservative 0;  
QY 3 CXXXXXC 9  
DB 31 CSRTAC 37  
RESULT 30  
D44007  
aptoxin IV - trap-door spider (Aptostichus schlingeri)  
N;Alternate names: insecticidal peptide Aps IV  
C;Species: Aptostichus schlingeri  
C;Date: 20-Aug-1999 #sequence\_revision 20-Aug-1999 #text\_change 09-Jul-2004  
C;Accession: D44007  
R;Skinner, W.S.; Dennis, P.A.; Li, J.P.; Quistad, G.B.  
Toxicon 30, 1043-1050, 1992  
A;Title: Identification of insecticidal peptides from venom of the trap-door spider, Ap  
A;Reference number: A44007; MUID:93069259; PMID:1440641  
A;Accession: D44007  
A;Molecule type: protein  
A;Residues: 1-76 <SKI>  
A;Cross-references: UNIPROT:P49269; PIDN:AA24050.1; PID:G259280  
A;Note: sequence extracted from NCBI backbone (NCBIP:119527)



C:Keywords: disulfide bond; toxin; venom

Query Match 31.5%; Score 17; DB 2; Length 76;  
Best Local Similarity 28.6%; Pred. NO. 7.3e+02;  
Matches 2; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 3 CXXXXXC 9  
Db 30 CSSSSEC 36

Search completed: April 27, 2005, 16:45:12  
Job time : 45 secs

**This Page Blank (uspto)**

**This Page Blank (uspto)**